

Revisiting the Nilpotent Polynomial Hales-Jewett Theorem

John H. Johnson and Florian Karl Richter
johnson.5316@osu.edu and richter.109@osu.edu

*Department of Mathematics
 The Ohio State University
 Columbus, Ohio*

July 20, 2016

Abstract

Answering a question posed by Bergelson and Leibman in [7], we establish a nilpotent version of the Polynomial Hales-Jewett Theorem that is a common generalization of both the main theorem in [7] and the Polynomial Hales-Jewett Theorem (PHJ).

Important to the formulation and the proof of our main theorem is Shuungula's, Zelenyuk's, and Zelenyuk's [30] notion of a relative syndetic set (relative with respect to two closed non-empty subsets of $\beta\mathbf{G}$). Using their language we first reformulate the classical PHJ as a theorem about abelian groups before we proceed to generalize it to nilpotent groups.

As a corollary of our main theorem we prove an extension of the restricted van der Waerden Theorem to nilpotent groups, which involves nilprogressions.

We also offer a reformulation of the Density Hales-Jewett Theorem based on an extension of the notion of relative syndeticity to a relative version of upper and lower Banach density.

Keywords: Polynomial Hales-Jewett Theorem; Ramsey theory; algebra in the Stone-Ćech compactification; nilpotent groups; nilprogressions; filters; syndetic sets

Contents

1 Introduction

2

2	Topological Algebra of Closed Subsemigroups of $\beta\mathbf{S}$	10
3	Connections between Closed Subsemigroups of $\beta\mathbf{S}$ and Hales-Jewett Type Theorems	12
4	Banach Density in Semigroups	15
5	Upper and Lower $(\mathcal{F}, \mathcal{G})$ -Density	17
6	Connections between Theorem F and the Density Hales-Jewett Theorem	21
7	A Proof of Theorem D	28
A	Appendix	37

1. Introduction

In [7, Remark 6.4] it was asked by Bergelson and Leibman if it is possible to formulate and prove a “full-fledged nilpotent Polynomial Hales-Jewett Theorem”, i.e., a theorem that complements their extension of the Polynomial van der Waerden Theorem to nilpotent groups and, at the same time, contains their celebrated Polynomial Hales-Jewett Theorem as a special case. In this paper we offer an affirmative answer to this question in the form of Theorem D.

We immediately give a paraphrased version below but the full formulation can be found on page 8.

Paraphrased Main Theorem. *Given any finite collection \mathbf{P} of “polynomial mappings” from the collection of all finite non-empty subsets of \mathbb{N} to a nilpotent group (G, \cdot) , there exist “restrictive conditions” on a subset of G such that any coloring of this subset contains monochromatic polynomial progressions of the form $\{aP(\alpha) : P \in \mathbf{P}\}$ for some $a \in G$ and some finite non-empty $\alpha \subset \mathbb{N}$.*

The notion of “polynomial mappings” is well established in polynomial Ramsey theory type results (see Definition 1.3 below). To precisely formulate Theorem D we therefore focus on motivating and explaining what we mean by “restrictive conditions”: which depends on the notions of syndeticity and filters.

Let (\mathbf{S}, \cdot) be a semigroup. A subset $A \subset \mathbf{S}$ is called **(right-)syndetic** if there exists a finite non-empty set $K \subset \mathbf{S}$ such that $K^{-1}A = \mathbf{S}$, where $K^{-1}A$

denotes the set $\{s \in \mathbf{S} : ks \in A \text{ for some } k \in K\}$. In many ways syndetic sets are ‘large’ enough to carry an abundance of semigroup structure. For instance Kakeya and Morimoto [27, Theorem I] observed, using van der Waerden’s theorem [33], that any syndetic set in \mathbb{N} contains arbitrarily long arithmetic progressions.

In [30] the authors Shuungula, Zelenyuk and Zelenyuk extended the notion of syndeticity: Given two filters¹ \mathcal{F} and \mathcal{G} on a semigroup \mathbf{S} , a set $A \subset \mathbf{S}$ is called **$(\mathcal{F}, \mathcal{G})$ -syndetic** if for every set $V \in \mathcal{F}$ there exists a finite non-empty set $K \subset V$ such that $K^{-1}A$ lies in \mathcal{G} . If $\mathcal{F} = \mathcal{G}$ then we say \mathcal{F} -syndetic instead of $(\mathcal{F}, \mathcal{G})$ -syndetic. Note, regular syndeticity corresponds to the special case where \mathcal{F} and \mathcal{G} are equal to the trivial filter on \mathbf{S} , i.e., $\mathcal{F} = \mathcal{G} = \{\mathbf{S}\}$.

Given two filters \mathcal{F} and \mathcal{G} on a semigroup (\mathbf{S}, \cdot) we define the **filter product** $\mathcal{F} \cdot \mathcal{G}$ according to the rule

$$A \in \mathcal{F} \cdot \mathcal{G} \iff \{x \in S : \{y \in S : x \cdot y \in A\} \in \mathcal{G}\} \in \mathcal{F}. \quad (1.1)$$

It can easily be checked that the filter product of two filters is itself a filter.

A filter \mathcal{F} is called **idempotent** if it satisfies $\mathcal{F} \cdot \mathcal{F} \supset \mathcal{F}$. The trivial filter $\{\mathbf{S}\}$ is an example of an idempotent filter. We give other examples below, see for instance the filter constructed in Example 1.1 or the filter given by Eq. (3.2). A special class of idempotent filters is the class of idempotent ultrafilters². We remark that idempotent ultrafilters have been studied extensively due to their applicability to Ramsey Theory and Ergodic Ramsey Theory (see [2, 3, 4, 8, 9, 19, 26, 32]).

Example 1.1. We use $\mathcal{P}_f(X)$ to denote the collection of all non-empty finite subsets of a set X . In particular, $\mathcal{P}_f(\mathbb{N})$ denotes the collection of all non-empty finite subsets of \mathbb{N} . For $\alpha, \beta \in \mathcal{P}_f(\mathbb{N})$ we write $\alpha < \beta$ if $\max \alpha < \min \beta$. Define

$$\mathbb{E} := \{V \subset \mathcal{P}_f(\mathbb{N}) : \exists \beta \text{ s.t. } \{\alpha : \alpha > \beta\} \subset V\}.$$

Then \mathbb{E} is an idempotent filter on the semigroup $(\mathcal{P}_f(\mathbb{N}), \cup)$. This filter can be thought of as the archetypical idempotent filter. Indeed, all explicit examples of idempotent filters that we give in this paper are variations of \mathbb{E} .

¹ If \mathcal{F} is a collection of subsets of a set X , then \mathcal{F} is called a **filter** on X if it satisfies

- $\emptyset \notin \mathcal{F}$ and $X \in \mathcal{F}$; (\mathcal{F} is proper)
- If $A \in \mathcal{F}$ and $B \supset A$ then $B \in \mathcal{F}$; (\mathcal{F} is upward closed)
- If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$. (\mathcal{F} has the finite intersection property)

²An **ultrafilter** is a maximal filter, i.e., a filter that is not properly contained in another filter.

It turns out that if \mathcal{F} is an idempotent filter, then \mathcal{F} -syndetic sets behave similarly to syndetic sets. In particular, if \mathcal{F} is idempotent then many theorems in Ramsey theory regarding syndetic sets can be generalized to \mathcal{F} -syndetic sets. Interestingly, such theorems are closely related to Hales-Jewett type theorems.

In order to further explore this analogy between syndetic sets and \mathcal{F} -syndetic sets, let us recall a generalization of van der Waerden's Theorem, which is due to Furstenberg and Weiss, known as the IP van der Waerden Theorem [22]. If (\mathbf{S}, \cdot) is a semigroup then a function $x : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ is called an **IP-set** if $x(\alpha \cup \beta) = x(\alpha) \cdot x(\beta)$ for all $\alpha, \beta \in \mathcal{P}_f(\mathbb{N})$ with $\alpha < \beta$.

IP van der Waerden Theorem (cf. [22, Section 3] and [18, Subsection 2.5]). *Let $k \in \mathbb{N}$, let $(\mathbf{S}, +)$ be a commutative semigroup and let $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ be IP-sets. Then for any syndetic set $A \subset \mathbf{S}$ there are $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{S}$ such that $\{a + x_1(\alpha), \dots, a + x_k(\alpha)\} \subset A$.*

Let us now see how the IP van der Waerden Theorem generalizes if we pass from syndetic sets to \mathcal{F} -syndetic sets. We need the following definition, where the terminology is based on the analogy between filters and measurable sets.

Definition 1.2. Let \mathbf{S} be a semigroup with identity $1_{\mathbf{S}}$, let \mathcal{F} be a filter on \mathbf{S} , let $x : \mathcal{P}_f(\mathbb{N}) \rightarrow X$ be a map and let \mathbb{E} be as in Example 1.1. We say that x is **\mathcal{F} -measurable** (or **$(\mathbb{E}, \mathcal{F})$ -measurable**) if for all $V \in \mathcal{F}$ the set $x^{-1}(V \cup \{1_{\mathbf{S}}\})$ is contained in \mathbb{E} .

Theorem A (IP van der Waerden Theorem for \mathcal{F} -syndetic sets). *Let $k \in \mathbb{N}$, let \mathcal{F} be an idempotent filter on a commutative semigroup $(\mathbf{S}, +)$ and let $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ be \mathcal{F} -measurable IP-sets. Suppose $A \subset \mathbf{S}$ is \mathcal{F} -syndetic. Then there are $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{S}$ such that $\{a + x_1(\alpha), \dots, a + x_k(\alpha)\} \subset A$.*

In Section 3 we prove that Theorem A implies the Hales-Jewett Theorem and from the method of this proof it will be evident that Theorem A and the Hales-Jewett Theorem are in fact equivalent (in the intuitive sense that one theorem is easily derivable from the other). For an independent proof of Theorem A we refer the reader to [16, Lemma 3.2] where a very similar result is proven.

In [5, Corollary 1.12], Bergelson and Leibman prove a polynomial extension of the IP van der Waerden Theorem and in [7] they further generalize it to nilpotent groups. We refer to their most general theorem as the Polynomial IP van der Waerden Theorem for Nilpotent Groups. For the statement

of this theorem we need the following definitions. Let (\mathbf{G}, \cdot) be a group. If $x, y : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ are two mappings then we say

$$x = y \quad \mathbb{E}\text{-a.e.} \quad \text{if and only if} \quad \{\alpha : x(\alpha) = y(\alpha)\} \in \mathbb{E}.$$

Also, given a map $x : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ and given $\beta \in \mathcal{P}_f(\mathbb{N})$ we define the discrete derivative $D_\beta x : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ as

$$D_\beta x(\alpha) = (x(\alpha))^{-1} x(\alpha \cup \beta) (x(\beta))^{-1}.$$

Definition 1.3 (Polynomial Mappings). A map $P : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ is called a **polynomial mapping of degree 1** if P is equal \mathbb{E} -a.e. to an IP-set. The map P is called a **polynomial mapping of degree d** if for \mathbb{E} -many³ β the mapping $D_\beta P$ is a polynomial of degree $d - 1$.

One can easily construct relevant examples of polynomial mappings in the following way: Suppose $x : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbb{N}$ is an IP-set and $p \in \mathbb{Z}[x]$ is a polynomial of degree d satisfying $p(0) = 0$. Then $P(\alpha) := p(x(\alpha))$ is a polynomial mapping of degree at most d . More involved examples of polynomial mappings – especially in the setting of nilpotent groups – are given in Example 7.1.

Polynomial IP van der Waerden Theorem for Nilpotent Groups ([7, Theorem 4.4]). *Let (\mathbf{G}, \cdot) be a nilpotent group, let $P_1, \dots, P_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ be polynomial mappings and let $A \subset \mathbf{G}$ be syndetic. Then there are $a \in \mathbf{G}$ and $\alpha \in \mathcal{P}_f(\mathbb{N})$ such that $\{aP_1(\alpha), \dots, aP_k(\alpha)\} \subset A$.*

Another strengthening of the IP van der Waerden Theorem is the celebrated Hales-Jewett Theorem [24]. It has also been generalized by Bergelson and Leibman [6] to include polynomial configurations and this extension is known as the Polynomial Hales-Jewett Theorem (PHJ).

PHJ has many equivalent forms (cf. [6, 7, 34]). To state one of them we need to define monomial mappings, which are a special type of polynomial mappings. We define monomial mappings only for abelian groups; an extension of this definition to more general groups (and semigroups) is not needed in this paper but can be found in [7, Subsection 1.3].

Definition 1.4 (Monomial Mappings). Let $(\mathbf{G}, +)$ be an abelian group. Given a map $u : \mathbb{N}^d \rightarrow \mathbf{G}$ and $\gamma \in \mathcal{P}_f(\mathbb{N}^d)$ we define $u(\gamma) := \sum_{s \in \gamma} u(s)$. A map $P : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ is called a **monomial mapping of degree d** if there

³The quantifier ‘for \mathcal{F} -many x in X ’ is defined in the following way: Given a filter \mathcal{F} on a set X , we say that property $P(x)$ holds for \mathcal{F} -many x if and only if the set $\{x \in X : P(x) \text{ holds}\}$ is contained in \mathcal{F} .

exists a map $u : \mathbb{N}^d \rightarrow \mathbf{G}$ such that P is \mathbb{E} -a.e. equal to the map $\alpha \mapsto u(\alpha^d)$. In this case we say that P is **induced** by u .

It is shown in [7] that for nilpotent groups (\mathbf{G}, \cdot) a map $P : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ is a polynomial mapping if and only if it can be written as product of monomial mappings (or sum of monomial mappings in the abelian case). Let us now state (a version of) the Polynomial Hales-Jewett Theorem.

Polynomial Hales-Jewett Theorem ([7, Theorem 6.5]). *Let $(\mathbf{G}, +)$ be an abelian group⁴. For any $k, d \in \mathbb{N}$ and any monomial mappings $P_1, \dots, P_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$, induced respectively by $u_1, \dots, u_k : \mathbb{N}^d \rightarrow \mathbf{G}$, and any finite coloring of \mathbf{G} there exist $\gamma_1, \dots, \gamma_k \in \mathcal{P}_f(\mathbb{N}^d)$ and $\alpha \in \mathcal{P}_f(\mathbb{N})$, with $\gamma_i \cap \alpha^d = \emptyset$, such that for $a = u_1(\gamma_1) + \dots + u_k(\gamma_k)$ the elements $a + P_1(\alpha), \dots, a + P_k(\alpha)$ have the same color.*

A particularly unappealing feature of this formulation of the Polynomial Hales-Jewett Theorem is the restriction on a , namely that $a = u_1(\gamma_1) + \dots + u_k(\gamma_k)$ with $\gamma_i \cap \alpha^d = \emptyset$. We can reformulate this condition using the notion of filter-syndetic sets as follows. Let $[N] := \{1, \dots, N\}$, let

$$U_N := \left\{ \sum_{i=1}^k u_i(\gamma_i) : \gamma_i \in \mathcal{P}_f(\mathbb{N}^d \setminus [N]^d), \gamma_i \cap \gamma_j = \emptyset, i \neq j \right\}$$

and let \mathcal{F} be defined as

$$\mathcal{F} = \mathcal{F}(u_1, \dots, u_k) := \{V \subset \mathbf{G} : \exists N \text{ s.t. } U_N \subset V\}. \quad (1.2)$$

It is straight forward to verify that \mathcal{F} is an idempotent filter on \mathbf{G} . This suggests that an appropriate replacement for the condition $a = u_1(\gamma_1) + \dots + u_k(\gamma_k)$ is given by the condition $\forall V \in \mathcal{F} \exists a \in V$. This leads to the following new reformulation of PHJ.

Theorem B. *Let $(\mathbf{G}, +)$ be an abelian group, let $k, d \in \mathbb{N}$ and let $P_1, \dots, P_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ be monomial mappings induced by $u_1, \dots, u_k : \mathbb{N}^d \rightarrow \mathbf{G}$. Let $\mathcal{F} = \mathcal{F}(u_1, \dots, u_k)$ be as in (1.2). Then for all \mathcal{F} -syndetic sets A there are $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{G}$ such that $\{a, a + P_1(\alpha), \dots, a + P_k(\alpha)\} \subset A$.*

It is not hard to show that Theorem B implies the Polynomial Hales-Jewett Theorem; we discuss this in more detail at the end of Section 3. The

⁴We remark that in this statement of the Polynomial Hales-Jewett Theorem one can replace $(\mathbf{G}, +)$ with any commutative semigroup $(\mathbf{S}, +)$. Note that any commutative semigroup is the homomorphic image of a cancellative commutative semigroup and any cancellative commutative semigroup can be embedded into an abelian group; this way one can easily derive the semigroup version from the group version.

reverse direction holds as well, however we don't prove this explicitly in this paper.

This raises the question: Can Theorem B be generalized to arbitrary polynomial mappings and arbitrary idempotent filters? Ideally, this generalization should contain Theorem A as a special case. To accomplish this goal we need one more definition.

Definition 1.5. Let \mathcal{F} be a filter on an abelian group \mathbf{G} with identity element $0_{\mathbf{G}}$ and let \mathbf{P} be a collection of polynomial mappings. We say that \mathbf{P} is a **good collection of \mathcal{F} -measurable polynomial mappings** if

- (i) $0_{\mathbf{G}} \in \mathbf{P}$ and every $P \in \mathbf{P}$ is \mathcal{F} -measurable;
- (ii) for all $R, P \in \mathbf{P}$ there exist \mathbb{E} -many β such that $R + D_{\beta}P \in \mathbf{P}$.

Theorem C (Polynomial van der Waerden Theorem for \mathcal{F} -syndetic sets). *Let \mathcal{F} be an idempotent filter on an abelian group $(\mathbf{G}, +)$ and let \mathbf{P} be a good collection of \mathcal{F} -measurable polynomial mappings. Then for all $P_1, \dots, P_k \in \mathbf{P}$ and all \mathcal{F} -syndetic sets A there are $\alpha \in \mathcal{P}_{\mathcal{F}}(\mathbb{N})$ and $a \in \mathbf{G}$ such that $\{a + P_1(\alpha), \dots, a + P_k(\alpha)\} \subset A$.*

Theorem A follows immediately from Theorem C. Also, Theorem B follows quickly from Theorem C after showing that monomial mappings induced by u_1, \dots, u_k (together with all their derivatives) are measurable with respect to $\mathcal{F} = \mathcal{F}(u_1, \dots, u_k)$. Subsequently, Theorem C also implies PHJ. As we will see later, one advantage of Theorem C over PHJ is that it is more amenable for generalizations to non-abelian groups.

In [7, Section 6] it is explained how the Polynomial IP van der Waerden Theorem for Nilpotent Groups lacks certain aspects of Hales-Jewett type theorems; in particular it does not contain PHJ as a special case (also see the remark after Theorem 3.3 in [35]). One of our main goals in this paper is to prove a generalization of Theorem C to nilpotent groups. Such a theorem would then be a common generalization of both the Polynomial IP van der Waerden Theorem for Nilpotent Groups and PHJ.

Let (\mathbf{G}, \cdot) be a nilpotent group. In the following, for $u, c \in \mathbf{G}$ we write u^c for the conjugate $c^{-1}uc$ and $[u, c]$ for the commutator $u^{-1}c^{-1}uc$.

Definition 1.6. Let \mathcal{F} be a filter on a nilpotent group \mathbf{G} with identity element $1_{\mathbf{G}}$ and let \mathbf{P} be a collection of polynomial mappings. We say that \mathbf{P} is a **good collection of \mathcal{F} -measurable polynomial mappings** if

- (I) $1_{\mathbf{G}} \in \mathbf{P}$ and every $P \in \mathbf{P}$ is \mathcal{F} -measurable (by abuse of language we use $1_{\mathbf{G}}$ to denote both the identity in \mathbf{G} and the constant polynomial mapping $\alpha \mapsto 1_{\mathbf{G}}$);
- (II) for all $R, P \in \mathbf{P}$ there are \mathcal{F} -many c such that $R^c[c, P] \in \mathbf{P}$;

(III) for all $R, P \in \mathbf{P}$ there are \mathcal{F} -many c and \mathbb{E} -many β such that $(RD_\beta P)^c \in \mathbf{P}$.

If \mathbf{G} is abelian then Definition 1.6 reduces to Definition 1.5. We remark that condition (I) is very natural in our setting whereas conditions (II) and (III) are technical necessities needed to perform a localized color focusing argument in Lemma 7.6. We refer the reader to Example 7.1 for explicit examples of good collections of \mathcal{F} -measurable polynomial mappings.

Theorem D (Nilpotent Polynomial van der Waerden Theorem for \mathcal{F} -syndetic sets). *Let \mathcal{F} be an idempotent filter on a nilpotent group (\mathbf{G}, \cdot) and let \mathbf{P} be a good collection of \mathcal{F} -measurable polynomial mappings. Then for all $P_1, \dots, P_k \in \mathbf{P}$ and all \mathcal{F} -syndetic sets A there are $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{G}$ such that $\{aP_1(\alpha), \dots, aP_k(\alpha)\} \subset A$.*

If \mathcal{F} is the trivial filter $\{\mathbf{G}\}$ then Theorem D reduces to the Polynomial IP van der Waerden Theorem for Nilpotent Groups. If \mathbf{G} is abelian then Theorem D reduces to Theorem C. If both \mathcal{F} is the trivial filter and \mathbf{G} is abelian then Theorem D becomes an analogue of [5, Corollary 1.12].

The Hales-Jewett Theorem and the Polynomial Hales-Jewett Theorem have found numerous applications to the combinatorics of abelian groups. For instance, it was shown by Spencer [31], using the Hales-Jewett Theorem, that there exists a set $V \subset \mathbb{N}$ containing no $k+1$ term arithmetic progressions and such that for any partition of V into finitely many classes, some class must contain a k -term arithmetic progression. This result is known as the restricted van der Waerden Theorem⁵.

Using Theorem D we can extend the restricted van der Waerden Theorem to nilpotent groups. In this extension the role of arithmetic progressions is taken over by nilprogressions. Nilprogressions⁶ are a well studied object that emerged from various generalizations of Freiman's theorem to non-abelian groups [11, 12, 13, 14]. For their definition let $\Sigma_{\leq k}$ denote the collection of all words $w(*_1, \dots, *_d)$ in the letters $*_1, \dots, *_d$ such that every letter $*_i$ appears at most $(k-1)$ times. Also, given a word $w(*_1, \dots, *_d)$ and elements x_1, \dots, x_d in a group (\mathbf{G}, \cdot) we use $w(x_1, \dots, x_d)$ to denote the group element of \mathbf{G} obtained by replacing all occurrences of the variable $*_i$ in the word $w(*_1, \dots, *_d)$ with x_i . Define a **nilprogression of step s , lenght k and**

⁵Note, a similar result was independently obtained by Nešetřil and Rödel [28] by directly proving a restricted version of van der Waerden's Theorem.

⁶We remark that in general nilprogressions depend on many different parameters. However, it will be convenient for us to replace the definition of nilprogressions, as it appears in the literature, with a slightly more simplified version that depends on fewer parameters.

rank d to be a set of the form

$$A := \{aw(x_1, \dots, x_d) : w \in \Sigma_{\leq k+1}\}$$

where a, x_1, \dots, x_d are elements in a s -step nilpotent group G . If $|A| = |\Sigma_{\leq k}|$ then we call A a **non-degenerated** nilprogression.

Theorem E (Restricted van der Waerden Theorem for nilprogressions). *For every $k \geq 1$ there exists a k -step nilpotent group (\mathbf{G}, \cdot) in two generators and a set $V \subset \mathbf{G}$ with the property that V does not contain any non-degenerated nilprogressions of step k , length $k+1$ and rank 2 but for any partition of V into finitely many classes, some class contains a non-degenerated nilprogressions of step k , length k and rank 2.*

We conjecture that analogues of Theorem E for nilprogressions of rank $d > 2$ also hold and can be derived from Theorem D, however we don't attempt to prove this conjecture in this paper. (Extending our current proof to prove this generalization seems to require constructing a k -step nilpotent group in d generators x_1, \dots, x_d where one can explicitly calculate all words $w(x_1, \dots, x_d) \in \Sigma_{\leq k+1}$.)

Throughout Sections 4–6 we advance from coloring Ramsey theory to density Ramsey theory. Among other things, we show that given a filter \mathcal{F} on a cancellative semigroup \mathbf{S} the notion of upper Banach density can be generalized to an appropriate analogue for what we call (upper) \mathcal{F} -density. Similar to the way syndetic sets and properties thereof generalize nicely to \mathcal{F} -syndetic sets, analogues of many of the properties of positive upper Banach density sets can also be proved for sets of positive \mathcal{F} -density. Moreover, many theorems from Ramsey theory concerning sets of positive Banach density, such as Furstenberg's and Katznelson's IP Szemerédi Theorem [19], can be generalized to sets with positive upper \mathcal{F} -density. For the precise definitions of Banach density and \mathcal{F} -density in semigroups, see Definition 4.1 and Definition 5.1.

To obtain the statement of the aforementioned IP Szemerédi Theorem one merely has to take the statement of the IP van der Waerden Theorem and replace 'syndetic' with 'positive upper Banach density'.

IP Szemerédi Theorem ([19]). *Let $k \in \mathbb{N}$, let $(\mathbf{S}, +)$ be a cancellative commutative semigroup and let $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ be IP-sets. Suppose $A \subset \mathbf{S}$ has positive upper Banach density. Then there are $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{S}$ such that $\{a + x_1(\alpha), \dots, a + x_k(\alpha)\}$ is contained in A .*

At this point it is easy to guess the right generalization of the the IP Szemerédi Theorem for \mathcal{F} -density sets: One merely needs to replace ' \mathcal{F} -

syndetic' with 'positive upper \mathcal{F} -density' in the statement of Theorem A and it is done.

Theorem F (Szemerédi's Theorem for \mathcal{F} -Density Sets). *Let $k \in \mathbb{N}$, let \mathcal{F} be an idempotent filter on a cancellative commutative semigroup $(\mathbf{S}, +)$ and let $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ be \mathcal{F} -measurable IP-sets. Suppose $A \subset \mathbf{S}$ has positive \mathcal{F} -density. Then there are $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{S}$ such that $\{a + x_1(\alpha), \dots, a + x_k(\alpha)\} \subset A$.*

In Section 6 we show that Theorem F is indeed equivalent to the Density Hales-Jewett Theorem [20, 21].

Finally, let us note that if we choose \mathcal{F} to be the trivial filter then the above theorems reduce to their antecedent counterparts. More precisely, if $\mathcal{F} = \{\mathbf{S}\}$, then Theorem A reduces to the IP van der Waerden Theorem whereas Theorem F reduces to the IP Szemerédi Theorem.

Acknowledgements: We thank Vitaly Bergelson for reading an early draft of this paper and providing several helpful comments and additional references, and also Joel Moreira for several helpful discussions.

2. Topological Algebra of Closed Subsemigroups of $\beta\mathbf{S}$

For a semigroup (\mathbf{S}, \cdot) , let $\beta\mathbf{S}$ denote the collection of all ultrafilters on \mathbf{S} . Note that $\beta\mathbf{S}$ is a semigroup when endowed with the operation given by equation (1.1). Given a subset $A \subset \mathbf{S}$ we define \overline{A} as the set $\{p \in \beta\mathbf{S} : A \in p\}$. It is well known that $\beta\mathbf{S}$ endowed with the topology generated by $\{\overline{A} : A \subset \mathbf{S}\}$ is a compact Hausdorff right-topological semigroup.⁷

There exists a natural one-to-one correspondence between filters on \mathbf{S} and non-empty closed subsets of $\beta\mathbf{S}$: If $T \subset \beta\mathbf{S}$ is non-empty and closed then the filter associated with T is defined as $\mathcal{F}_T := \{A \subset \mathbf{S} : T \subset \overline{A}\}$. Vice versa, if \mathcal{F} is a filter on \mathbf{S} then the **closure of the filter \mathcal{F}** , defined as the set $\overline{\mathcal{F}} := \bigcap_{A \in \mathcal{F}} \overline{A}$, is a non-empty and closed subset of $\beta\mathbf{S}$. Clearly, the closure of the filter \mathcal{F}_T is T and the filter associated with $\overline{\mathcal{F}}$ is \mathcal{F} .

We are particularly interested in idempotent filters. If \mathcal{F} is idempotent then the closure $T = \overline{\mathcal{F}}$ of \mathcal{F} is a closed subsemigroup of $\beta\mathbf{S}$. Note that the reverse is not true; there are closed subsemigroups whose corresponding filter is not idempotent.⁸ Also, we remark that in Theorems A, C, D and F the condition for \mathcal{F} to be idempotent can be relaxed to the condition for

⁷See [26] for a comprehensive discussion on the topological and algebraical aspects of $\beta\mathbf{S}$.

⁸See for instance [15] for a complete combinatorial characterization of closed subsemigroups of $\beta\mathbf{S}$.

$\overline{\mathcal{F}}$ to be a subsemigroup of $\beta\mathbf{S}$. Hence understanding closed subsemigroups of $\beta\mathbf{S}$ lies at the heart of understanding those theorems. Still, in this paper we will not push these theorems to such generality, since the authors were not able to derive any additional combinatorial applications from it.

Next, let us address ideals and idempotent elements of closed subsemigroups of $\beta\mathbf{S}$. Any compact Hausdorff right topological semigroup T has a smallest two sided ideal $K(T)$, which is the union of all minimal right ideals and is the union of all minimal left ideals (see [26]). Also, by the Ellis-Numakura theorem ([17, 29]), every compact Hausdorff right topological semigroup T contains at least one idempotent element. We denote the collection of all idempotent elements in T by $E(T)$.

Definition 2.1. Let \mathcal{F} and \mathcal{G} be filters on \mathbf{S} . A set $A \subset \mathbf{S}$ is called $(\mathcal{F}, \mathcal{G})$ -**thick**, if there exists a \mathcal{F} -large set V such that for all \mathcal{G} -large sets W and for all non-empty finite subsets $F \subset V$ there exists some $s \in W$ such that Fs is contained in A . A set $A \subset \mathbf{G}$ is called **piecewise $(\mathcal{F}, \mathcal{G})$ -syndetic** if A can be written as $A = B \cap C$ for some $(\mathcal{F}, \mathcal{G})$ -syndetic set B and some $(\mathcal{F}, \mathcal{G})$ -thick set C .

If $\mathcal{G} = \mathcal{F}$, then we say piecewise \mathcal{F} -syndetic and \mathcal{F} -thick instead of piecewise $(\mathcal{F}, \mathcal{G})$ -syndetic and $(\mathcal{F}, \mathcal{G})$ thick. Moreover, if $\mathcal{F} = \mathcal{G} = \{\mathbf{S}\}$, then we simply say syndetic, piecewise syndetic and thick⁹.

It follows immediately that a set is $(\mathcal{G}, \mathcal{F})$ -syndetic if and only if its complement is not $(\mathcal{G}, \mathcal{F})$ -thick.

Before we finish this section, let us provide the reader with algebraic characterizations of \mathcal{F} -syndetic sets and piecewise \mathcal{F} -syndetic sets.

Theorem 2.2 (Theorem 2.2, [30]). *Let $T = \overline{\mathcal{F}}$ be a closed subsemigroup of $\beta\mathbf{S}$ and let $p \in K(T)$. Then for any set $A \in p$ the set $A/p = \{x : x^{-1}A \in p\}$ is \mathcal{F} -syndetic.*

Theorem 2.3 (Theorem 2.3, [30]). *Let $T = \overline{\mathcal{F}}$ be a closed subsemigroup of $\beta\mathbf{S}$ and let $A \subset \mathbf{S}$. Then A is piecewise \mathcal{F} -syndetic if and only if $A \in p$ for some $p \in K(T)$.*

Remark 2.4. As an immediate consequence of Theorem 2.3 it follows that if A is piecewise \mathcal{F} -syndetic and A is partitioned into finitely many classes then at least one of the classes is piecewise \mathcal{F} -syndetic.

Remark 2.5. Due to Theorem 2.2 and Theorem 2.3 it is clear that in

⁹In non-abelian groups, one often distinguishes between right and left syndetic sets as well as between left and right thick sets. In fact, what we refer to as syndetic is also known as right syndetic, and what we refer to as thick is also known as left thick.

Theorems A, C and D we are allowed to replace ‘ \mathcal{F} -syndetic’ with ‘piecewise \mathcal{F} -syndetic’ and the statements remain true.

3. Connections between Closed Subsemigroups of $\beta\mathbf{S}$ and Hales-Jewett Type Theorems

In this section we discuss the Hales-Jewett Theorem (HJ) and its connections to closed subsemigroups of $\beta\mathbf{S}$. Let $k \in \mathbb{N}$ and let $[k]$ denote the set $\{1, 2, \dots, k\}$. We refer to elements in $[k]^n$ as words of length n in the letters $1, 2, \dots, k$. For $a \in [k]^n$ and $j \in [n]$ we usually write a_j to denote the j -th letter of the word a .

A **variable word** $w(*)$ is a word in the letters $[k] \cup \{*\}$ in which the letter $*$ (*the variable*) appears at least once. Given a variable word $w(*)$ we define the words $w(1), \dots, w(k)$ as the words that are obtained by replacing all instances of the letter $*$ in $w(*)$ with the letters $1, \dots, k$ respectively. The k -tuple $(w(1), \dots, w(k))$ is often referred to as a **combinatorial line**.

Finally, given two words $a \in [k]^n$ and $b \in [k]^m$, we define the concatenation of a and b , denoted by $a \frown b$, as the word in $[k]^{n+m}$ given by $(a \frown b)_j = a_j$ for $j \in [n]$ and $(a \frown b)_j = b_{j-n}$ for $j \in (n, n+m]$.

Using the above definitions, we can now state the Hales-Jewett Theorem.

Hales-Jewett Theorem ([24]). *Suppose $\bigcup_{n \in \mathbb{N}} [k]^n$ is colored into finitely many colors. Then there exists a monochromatic combinatorial line.*

Our next goal is to show how one can use Theorem A to derive the Hales-Jewett Theorem.

Proposition 3.1. *Theorem A implies the Hales-Jewett Theorem.*

Proof. Let $k \in \mathbb{N}$ be fixed. The idea is to choose a convenient idempotent filter \mathcal{F} on $(\mathbb{Z}, +)$ and suitable IP-sets $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbb{Z}$ such that HJ follows from Theorem A with the help of Theorem 2.2.

Let (x_j) be an arbitrary and sparse sequence of positive integers. Let

$$V_{k,N} := \left\{ \sum_{j=N}^n e_j x_j : e_j \in [0, k], n \geq N \right\}, \quad (3.1)$$

let

$$\mathcal{F} = \{V \subset \mathbb{N} : \exists N \in \mathbb{N} \text{ s.t. } V_{k,N} \subset V\} \quad (3.2)$$

and, for $1 \leq i \leq k$, let $x_i : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbb{N}$ be given by $x_i(\alpha) = (i-1) \sum_{j \in \alpha} x_j$. It is straight-forward to verify that \mathcal{F} is an idempotent filter and that x_i , $1 \leq i \leq k$, is a collection of \mathcal{F} -measurable IP-sets.

Let W_1 denote the collection of all words in $\bigcup_{n \in \mathbb{N}} [k]^n$ that do not end with the letter 1. There is an obvious map φ between W_1 and $V_{k,1}$ given by $\varphi(a_1 a_2 \cdots a_n) = \sum_{j=1}^n e_j x_j$ where $a_j = e_j + 1$. Note that if the sequence (x_j) was chosen sparse enough (for instance, choosing $x_j = k^j$ suffices), then φ is guaranteed to be bijective, because we excluded all words that end in the letter 1.

Let $T = \overline{\mathcal{F}}$ and let $p \in K(T)$ be arbitrary. Suppose $\bigcup_{n \in \mathbb{N}} [k]^n$ is colored into finitely many colors and assume this coloring corresponds to the partition $\bigcup_{n \in \mathbb{N}} [k]^n = \bigcup_{i=1}^r C_i$. It follows that $V_{k,1} = \bigcup_{i=1}^r \varphi^{-1}(C_i)$ and since $V_{k,1} \in \mathcal{F} \subset p$ we deduce that there exists some $i_0 \in [r]$ such that $\varphi^{-1}(C_{i_0})$ is contained in p .

According to Theorem 2.2 the set $A = \{x : \varphi^{-1}(C_{i_0}) - x \in p\}$ is \mathcal{F} -syndetic. Hence, by Theorem A, there are $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbb{N}$ such that $\{a + x_1(\alpha), \dots, a + x_k(\alpha)\} \subset A$. This implies that the set

$$\bigcap_{i=1}^k \varphi^{-1}(C_{i_0}) - (a + x_i(\alpha)) \in p$$

and therefore there exists a set $B \in p$ such that

$$\{a + b + x_i(\alpha) : b \in B, i \in [k]\} \subset \varphi^{-1}(C_{i_0}).$$

Since $a + b + x_i(\alpha) \in V_{k,1}$ it can be written as $a + b + x_i(\alpha) = \sum_{j=1}^m e_j x_j$ for some $m \geq 1$ and where $e_j = i$ whenever $j \in \alpha$. Finally, if $n > \max \alpha$ then the pre-image of the k -term arithmetic progression $a + b + x_1(\alpha), \dots, a + b + x_k(\alpha)$ under φ is a monochromatic combinatorial line. However, since we can pick b from an infinite set B we can pick b such that $n > \max \alpha$ is guaranteed. \square

Remark 3.2. One can use the filter \mathcal{F} constructed in the proof of Proposition 3.1 to quickly derive a proof of the restricted van der Waerden theorem: Take \mathcal{F} and $V_{k,1}$ to be defined as above. Certainly, $V_{k,1}$ does not contain any $k + 1$ -term arithmetic progressions. However, for any coloring of $V_{k,1}$ one of the colors will be piecewise \mathcal{F} -syndetic and therefore, using Theorem A and Remark 2.5, we see that this color contains a k -term arithmetic progression (as is shown in the proof of Proposition 3.1).

The polynomial counterpart to HJ is PHJ, whereas the polynomial counterpart of Theorem A is Theorem C. Thus, the following is an analogue of Proposition 3.1.

Proposition 3.3. *Theorem B – which is a special case of Theorem C – implies the Polynomial Hales-Jewett Theorem.*

Proof. Let $(\mathbf{G}, +)$ be an abelian group, let $k, d \in \mathbb{N}$ be arbitrary and let $P_1, \dots, P_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ be monomial mappings induced by $u_1, \dots, u_k : \mathbb{N}^d \rightarrow \mathbf{G}$ respectively. We have to show that for any finite coloring of \mathbf{G} there exist $\gamma_1, \dots, \gamma_k \in \mathcal{P}_f(\mathbb{N}^d)$ and $\alpha \in \mathcal{P}_f(\mathbb{N})$, with $\gamma_i \cap \alpha^d = \emptyset$, such that for $a = u_1(\gamma_1) + \dots + u_k(\gamma_k)$ the elements $a + P_1(\alpha), \dots, a + P_k(\alpha)$ have the same color.

First, we claim that without loss of generality we can assume that \mathbf{G} is the free abelian group in $\mathbb{N}^d \times [k]$ generators, which we denote by $(e_{n,i})_{n \in \mathbb{N}^d, i \in [k]}$, and $u_i(n) := e_{n,i}$. To verify this claim observe that given any other abelian group G' and any other k -tuple of maps $u'_1, \dots, u'_k : \mathbb{N}^d \rightarrow G'$ there exists an obvious homomorphism $\Phi : G \rightarrow G'$ such that $u'_i = \Phi \circ u_i$, namely the homomorphism uniquely determined by $\Phi(e_{n,i}) = u'_i(n)$. This implies that $P'_i = \Phi \circ P_i$, where P'_i denotes the monomial induced by u'_i . Hence, for any finite coloring of G' the pull-back under Φ yields a finite coloring of G and if $\gamma_1, \dots, \gamma_k \in \mathcal{P}_f(\mathbb{N}^d)$ and $\alpha \in \mathcal{P}_f(\mathbb{N})$, with $\gamma_i \cap \alpha^d = \emptyset$, are found such that $a = u_1(\gamma_1) + \dots + u_k(\gamma_k)$ and the elements $a + P_1(\alpha), \dots, a + P_k(\alpha)$ have the same color, then $\Phi(a) = u'_1(\gamma_1) + \dots + u'_k(\gamma_k)$ and the elements $\Phi(a) + P'_1(\alpha), \dots, \Phi(a) + P'_k(\alpha)$ also have the same color.

Now suppose we are given an arbitrary finite coloring of \mathbf{G} . For convenience we view this finite coloring as a finite partition $\mathbf{G} = \bigcup_{i=1}^r C_i$. Let

$$U_N := \left\{ \sum_{i=1}^k \sum_{n \in \gamma_i} e_{n,i} : \gamma_i \in \mathcal{P}_f(\mathbb{N}^d \setminus [N]^d), \gamma_i \cap \gamma_j = \emptyset \text{ for } i \neq j \right\}$$

and let \mathcal{F} be defined as

$$\mathcal{F} := \{V \subset \mathbf{S} : \exists N \text{ s.t. } U_N \subset V\}.$$

It is straightforward to check that \mathcal{F} is an idempotent filter; in fact, \mathcal{F} coincides with the filter construed in (1.2).

Let $T := \overline{\mathcal{F}}$ denote the closed subsemigroup of $\beta\mathbf{G}$ corresponding to \mathcal{F} and let $p \in K(T)$ be arbitrary. Using the ultrafilter property of p we can find $i_0 \in \{1, \dots, r\}$ such that $C_{i_0} \in p$. Therefore the set $D := C_{i_0} \cap U_1$ is also contained in p . Let A denote the set $D - p := \{x \in \mathbf{G} : -x + D \in p\}$ and observe that by Theorem 2.2 the set A is \mathcal{F} -syndetic.

Now we can apply Theorem B. Hence, we can find $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $c \in \mathbf{G}$ such that $\{c, c + P_1(\alpha), \dots, c + P_k(\alpha)\} \subset A$. Since $A = D - p$ we deduce that there exists a group element $b \in \mathbf{G}$ such that if we put $a := b + c$ then $\{a, a + P_1(\alpha), \dots, a + P_k(\alpha)\} \subset D$. In particular all the elements $a, a + P_1(\alpha), \dots, a + P_k(\alpha)$ have the same color. Finally, by examining the

construction of U_1 we deduce that $\{a, a + P_1(\alpha), \dots, a + P_k(\alpha)\} \subset U_1$ holds only if a is of the form $a = u_1(\gamma_1) + \dots + u_k(\gamma_k)$ with $\gamma_i \cap \alpha^d = \emptyset$. \square

4. Banach Density in Semigroups

Let (\mathbf{S}, \cdot) be a countable semigroup. We say that \mathbf{S} satisfies the **strong Følner condition** (SFC) if

$$\forall H \in \mathcal{P}_f(\mathbf{S}), \forall \varepsilon > 0, \exists F \in \mathcal{P}_f(\mathbf{S}), \forall s \in H, \frac{|F \triangle sF|}{|F|} < \varepsilon. \quad (\text{SFC})$$

We call a sequence of non-empty finite subsets (F_N) of \mathbf{S} a (left) Følner-sequence if

$$\lim_{N \rightarrow \infty} \frac{|F_N \triangle sF_N|}{|F_N|} = 0, \quad \forall s \in \mathbf{S}.$$

Note that (SFC) implies the existence of a (left) Følner sequence and also implies (but is not equivalent to) the existence of a left invariant mean on \mathbf{S} . To us it will also be important that $|F \triangle sF| < \varepsilon|F|$ implies $|s^{-1}F \cap F| \geq (1 - \varepsilon)|F|$ ¹⁰. For a more detailed discussion on Følner conditions, Følner sequences and invariant means on discrete semigroups see [25].

Definition 4.1. The **upper and lower Banach density** of a set $A \subset \mathbf{S}$, denoted by $d^*(A)$ and $d_*(A)$ respectively, are defined as

$$d^*(A) = \sup_{(F_N)} \limsup_{N \rightarrow \infty} \frac{|A \cap F_N|}{|F_N|} \quad \text{and} \quad d_*(A) = \inf_{(F_N)} \liminf_{N \rightarrow \infty} \frac{|A \cap F_N|}{|F_N|},$$

where the supremum and infimum are taken over all (left) Følner-sequences (F_N) of \mathbf{S} .

Remark 4.2. Let \mathbf{S} be a cancellative semigroup satisfying (SFC). Then d^* and d_* exhibit the following well-known properties:

- (B1) Both d^* and d_* are well defined, subadditive and satisfy $d^*(t^{-1}A) \geq d^*(A)$ and $d_*(t^{-1}A) \geq d_*(A)$ for all $t \in \mathbf{S}$.
- (B2) A set $A \subset \mathbf{S}$ is thick if and only if $d^*(A) = 1$.
- (B3) A set $A \subset \mathbf{S}$ is syndetic if and only if $d_*(A) > 0$.
- (B4) Partition regularity: If $d^*(A) > 0$ and $A = A_1 \cup A_2$ then either $d^*(A_1) > 0$ or $d^*(A_2) > 0$.

¹⁰This follows from $|F| - |s^{-1}F \cap F| \leq |F \cup sF| - |s^{-1}F \cap F| \leq |F \cup sF| - |F \cap sF| = |F \triangle sF| \leq \varepsilon|F|$.

(B5) Khintchine's recurrence theorem: If $d^*(A) > 0$ then the set $\{t : d^*(A \cap t^{-1}A) > 0\}$ is syndetic.

We end this section with another characterization of upper Banach density. In many ways, sets of positive upper Banach density can be thought of as rarified thick sets. To better understand this analogy, let us recall that a set is thick if and only if for every finite non-empty set $F \subset \mathbf{S}$ there exists some $s \in \mathbf{S}$ such that all elements of Fs – or one might say $|F|$ many elements of Fs – are contained in A . The next theorem, which holds in any cancellative semigroup, asserts that the upper Banach density of a set A is at least δ if and only if for every finite set F there exists some $s \in \mathbf{S}$ such that at least $\delta|F|$ many elements of Fs are contained in A . This interpretation of d^* was also observed by Griesmer in [23, Corollary 9.2] in the context of the positive integers.

Theorem G. *Suppose \mathbf{S} is a cancellative semigroup satisfying (SFC). Let A be a non-empty subset of \mathbf{S} . Then*

$$d^*(A) = \sup \left\{ \tau \in [0, 1) : \forall F \in \mathcal{P}_f(\mathbf{S}), \exists s \in \mathbf{S}, \text{ s.t. } |A \cap Fs| > \tau|F| \right\}.$$

Proof. First, suppose that for every finite non-empty set $F \subset \mathbf{S}$ there exists $s \in \mathbf{S}$ such that

$$|A \cap Fs| > \tau|F|. \quad (4.1)$$

From this we will derive that $d^*(A) \geq \tau$. Let (F_N) be an arbitrary Følner sequence. Choosing $F = F_N$ in (4.1) we get that $|A \cap F_N t_N| > \tau|F_N| \geq \tau|F_N t_N|$ for some $t_N \in \mathbf{S}$. However, since \mathbf{S} is cancellative, $(F_N t_N)$ is also a left Følner sequence. This shows that $d^*(A) \geq \tau$.

Next, assume that $d^*(A) > \tau$. From this we want to derive that for every finite set $F \subset \mathbf{S}$ there exists $s \in \mathbf{S}$ such that $|A \cap Fs| > \tau|F|$. Let us choose a Følner sequence (F_N) such that $\lim_{N \rightarrow \infty} \frac{|A \cap F_N|}{|F_N|} > \tau$. For N sufficiently large we have that $|t^{-1}A \cap F_N| > \tau|F_N|$ for all $t \in F$. Then, by applying Lemma A.2 to the sequence $(t^{-1}A)_{t \in F}$, we can find $Y \subset F$ with $|Y| > \tau|F|$ such that $\bigcap_{t \in Y} t^{-1}A \neq \emptyset$. Let s be any point in this intersection. Then

$$Ys \subset A \cap Fs,$$

which implies $|A \cap Fs| > \tau|F|$. □

5. Upper and Lower $(\mathcal{F}, \mathcal{G})$ -Density

Let us now give a formal definition of upper and lower $(\mathcal{F}, \mathcal{G})$ -density. This definition is inspired by Theorem G. Assume \mathcal{F} and \mathcal{G} are filters on a countable semigroup \mathbf{S} .

Definition 5.1. For $A \subset \mathbf{S}$ and $\delta \in [0, 1]$ we say the **upper $(\mathcal{F}, \mathcal{G})$ -density** of A is at least δ , and write $d^*(A; \mathcal{F}, \mathcal{G}) \geq \delta$, if

$$\exists V \in \mathcal{F}, \forall \tau \in [0, \delta), \forall F \in \mathcal{P}_f(V), \forall W \in \mathcal{G}, \exists s \in W, |A \cap Fs| > \tau|F|.$$

In the following, we will write $d^*(A; \mathcal{F}, \mathcal{G}) > 0$ if we mean that there exists some $\delta \in (0, 1]$ such that $d^*(A; \mathcal{F}, \mathcal{G}) \geq \delta$. If $\mathcal{F} = \mathcal{G}$ then we say upper \mathcal{F} -density instead of upper $(\mathcal{F}, \mathcal{G})$ -density and we write $d^*(A; \mathcal{F})$ instead of $d^*(A; \mathcal{G}, \mathcal{F})$.

Example 5.2. If p and q are ultrafilters on \mathbb{N} , then the notions (p, q) -thick, (p, q) -syndetic and positive upper (p, q) -density coincide. As a matter of fact, A is any of the above if and only if $A \in p + q$.

Example 5.3. Suppose \mathbf{G} is an amenable group. If $\mathcal{F} = \mathcal{G} = \{\mathbf{G}\}$, then $d^*(A; \mathcal{F}, \mathcal{G}) \geq \delta$ if and only if $d^*(A) \geq \delta$. This follows immediately from Theorem G.

We can also define the **lower $(\mathcal{F}, \mathcal{G})$ -density**: We say that $d_*(A; \mathcal{F}, \mathcal{G}) > \delta$, if

$$\forall V \in \mathcal{F}, \exists \tau \in (\delta, 1], \exists F \in \mathcal{P}_f(V), \{t : |A \cap Ft| \geq \tau|F|\} \in \mathcal{G},$$

and we say that $d_*(A; \mathcal{F}, \mathcal{G}) = 1$ if $d_*(A; \mathcal{F}, \mathcal{G}) > \delta$ for all $\delta \in [0, 1)$.

Next, we would like to derive analogues of the properties listed in Remark 4.2 for upper and lower $(\mathcal{F}, \mathcal{G})$ -density. For this, we need one more definition. We say a pair of filters $(\mathcal{F}, \mathcal{G})$ satisfies the **Følner condition (FC)** if for all $\varepsilon > 0$ we have

$$\forall W \in \mathcal{F}, \exists V \in \mathcal{G}, \forall H \in \mathcal{P}_f(V), \exists F \in \mathcal{P}_f(W), \forall s \in H, \frac{|F \Delta sF|}{|F|} < \varepsilon. \quad (\text{FC})$$

Example 5.4. Let (x_n) be any sequence of natural numbers and define

$$U_N := \left\{ \sum_{j=N}^n e_j x_j : e_j \in [0, 2^j), n \geq N \right\}. \quad (5.1)$$

The sets U_N can be thought of as a finite-sums-set with increasing multiplicities. Sets of this kind, also known as $\overline{\mathbb{IP}}$ -sets, have first been studied in [10]. Let \mathcal{G} be the filter $\{A \subset \mathbb{Z} : \exists N \in \mathbb{N} \text{ such that } U_N \subset A\}$. Let $k \in \mathbb{N}$ be arbitrary, let $V_{k,N}$ be as in equation (3.1) and let \mathcal{F} be the filter $\{A \subset \mathbb{Z} : \exists N \in \mathbb{N} \text{ such that } V_{k,N} \subset A\}$. Then a straight forward calculation shows that the pair $(\mathcal{F}, \mathcal{G})$ satisfies (FC).

Theorem 5.5. *Suppose \mathcal{F} and \mathcal{G} are filters on a cancellative semigroup \mathbf{S} .*

- (F-B1) *Assume \mathcal{F} is idempotent. If $d^*(A; \mathcal{F}, \mathcal{G}) \geq \delta$ then $d^*(t^{-1}A; \mathcal{F}, \mathcal{G}) \geq \delta$ for \mathcal{F} -many t . Likewise, if $d_*(A; \mathcal{F}, \mathcal{G}) \geq \delta$ then $d_*(t^{-1}A; \mathcal{F}, \mathcal{G}) \geq \delta$ for \mathcal{F} -many t .*
- (F-B2) *A set $A \subset \mathbf{S}$ is $(\mathcal{F}, \mathcal{G})$ -thick if and only if $d^*(A; \mathcal{F}, \mathcal{G}) = 1$.*
- (F-B3) *A set $A \subset \mathbf{S}$ is $(\mathcal{F}, \mathcal{G})$ -syndetic if and only if $d_*(A; \mathcal{F}, \mathcal{G}) > 0$.*
- (F-B4) *Partition regularity: Suppose the pair $(\mathcal{G}, \mathcal{F})$ satisfies condition (FC). If A has positive upper $(\mathcal{F}, \{\mathbf{S}\})$ -density and $A = A_1 \cup A_2$ then either A_1 or A_2 has positive upper $(\mathcal{G}, \{\mathbf{S}\})$ -density.*
- (F-B5) *An analogue of Khintchine's recurrence theorem: Suppose \mathcal{H} is another filter and the pairs $(\mathcal{H}, \mathcal{G})$ and $(\mathcal{G}, \mathcal{F})$ satisfy condition (FC). If $d^*(A; \mathcal{F}, \{\mathbf{S}\}) > 0$ then the set $\{t \in \mathbf{S} : d^*(A \cap t^{-1}A; \mathcal{H}, \{\mathbf{S}\}) > 0\}$ is \mathcal{G} -syndetic.*

Proof of (F-B1). Assume $d^*(A; \mathcal{F}, \mathcal{G}) \geq \delta$. This means there exists $V \in \mathcal{F}$ such that for all $\tau \in [0, \delta)$, all $W \in \mathcal{G}$ and all $F \in \mathcal{P}_f(V)$ we can find $t \in W$ such that $|A \cap Ft| > \tau|F|$. Since \mathcal{F} is idempotent, the set $V/\mathcal{F} := \{s : s^{-1}V \in \mathcal{F}\}$ is in \mathcal{F} . We claim that $d^*(s^{-1}A; \mathcal{F}, \mathcal{G}) \geq \delta$ for all $s \in V/\mathcal{F}$. Let $s \in V/\mathcal{F}$ be arbitrary and define $\tilde{V} := V \cap s^{-1}V$. Clearly $\tilde{V} \in \mathcal{F}$. Let $\tau \in [0, \delta)$, $W \in \mathcal{G}$ and $F \in \mathcal{P}_f(\tilde{V})$ be arbitrary. Note that $sF \subset V$ and therefore we can find $t \in W$ such that $|A \cap sFt| > \tau|sF| = \tau|F|$. This implies that $|s^{-1}A \cap Ft| \geq |A \cap sFt| \geq \tau|F|$. Therefore, $d^*(s^{-1}A; \mathcal{F}, \mathcal{G}) \geq \delta$ for all $s \in V/\mathcal{F}$.

Regarding the second part of (F-B1) let us simply remark that the proof of $d_*(s^{-1}A; \mathcal{F}, \mathcal{G}) \geq \delta$ for \mathcal{F} -many s is analogous to the proof of the first part of (F-B1) and is therefore omitted. \square

Proof of (F-B2). The forwards direction follows immediately from the definition of $(\mathcal{F}, \mathcal{G})$ -thickness. For the backwards direction assume that there is $V \in \mathcal{F}$ such that for all $\tau \in [0, 1)$, all $W \in \mathcal{G}$ and all $F \in \mathcal{P}_f(V)$ there exists $t \in V$ with $|A \cap Ft| > \tau|F|$. Let $F \in \mathcal{P}_f(V)$ be arbitrary. Take any $\tau \in [0, 1)$ with $\tau > 1 - \frac{1}{|F|}$ and find $t \in W$ as guaranteed. Then, since

$$\frac{|A \cap Ft|}{|F|} > 1 - \frac{1}{|F|},$$

it follows that $Ft \subset A$. Indeed this shows that A is $(\mathcal{F}, \mathcal{G})$ -thick. \square

Proof of (F-B3). This easily follows from (F-B2): A set A is $(\mathcal{F}, \mathcal{G})$ -syndetic if and only if its complement is not $(\mathcal{F}, \mathcal{G})$ -thick. However, according to part (F-B2), this is the case if and only if the complement of A does not have upper $(\mathcal{F}, \mathcal{G})$ -density equal to 1. Using the inequality

$$d^*(A; \mathcal{F}, \mathcal{G}) + d_*(\mathbf{S} \setminus A; \mathcal{F}, \mathcal{G}) \geq 1,$$

we see that this is equivalent to A having positive lower $(\mathcal{F}, \mathcal{G})$ -density. \square

Proof of (F-B4). Suppose $A = A_1 \cup A_2$ with $d^*(A; \mathcal{F}, \{\mathbf{S}\}) > 0$. This means there exists $\tau > 0$ and $W \in \mathcal{F}$ such that for all $F \in \mathcal{P}_f(W)$ we can find $t \in \mathbf{S}$ such that $|A \cap Ft| > \tau|F|$. Then, using (FC), we can find $V \in \mathcal{G}$ such that for all $H \in \mathcal{P}_f(V)$ we can find a non-empty finite subset $F \subset W$ such that $|F \triangle sF| < \varepsilon|F|$ for all $s \in H$. Let $H_1 \subset H_2 \subset \dots$ be a nested sequence of finite subsets of V with the property that $\bigcup_{n \in \mathbb{N}} H_n = V$. Then for any such H_n we can find a non-empty finite set $F_n \subset W$ such that $|F_n \triangle sF_n| < \frac{\tau}{4}|F_n|$ for all $s \in H_n$. Also, since $F_n \in \mathcal{P}_f(W)$, we can find t_n such that $|A \cap F_n t_n| > \tau|F_n|$.

In particular, this means that either $|A_1 \cap F_n t_n| > \frac{\tau}{2}|F_n|$ or $|A_2 \cap F_n t_n| > \frac{\tau}{2}|F_n|$. Therefore, for any $n \in \mathbb{N}$, there exists $i_n \in \{1, 2\}$ such that $|A_{i_n} \cap F_n t_n| > \frac{\tau}{2}|F_n|$. Certainly, either the set $\{n : i_n = 1\}$ or the set $\{n : i_n = 2\}$ is infinite. Let us assume without loss of generality that the former set is infinite. We claim that this implies that A_1 has positive upper $(\mathcal{G}, \{\mathbf{S}\})$ -density. More precisely, we claim that for all $H \in \mathcal{P}_f(V)$ there exists $r \in \mathbf{S}$ such that $|A_1 \cap Hr| > \frac{\tau}{4}|H|$. Let $H \in \mathcal{P}_f(V)$ be arbitrary. Pick $n \in \mathbb{N}$ such that $H_n \supset H$ and thereafter pick $n' \geq n$ such that $i_{n'} = 1$. Put $t = t_{n'}$. We get that $|A_1 \cap F_{n'} t| > \frac{\tau}{2}|F_{n'}|$ and moreover, since $|F_{n'} \triangle sF_{n'}| < \frac{\tau}{4}|F_{n'}|$ for all $s \in H$, we get that $|A_1 \cap sF_{n'} t| > \frac{\tau}{4}|F_{n'}|$ for all $s \in H$. From this it follows that $|s^{-1}A_1/t \cap F_{n'}| > \frac{\tau}{4}|F_{n'}|$ for all $s \in H$, where $s^{-1}A_1/t$ means the set $\{x \in \mathbf{S} : sxt \in A_1\}$. Then, by applying Lemma A.2 to the sequence $(s^{-1}A_1/t)_{s \in H}$, we can find $Y \subset H$ with $|Y| > \frac{\tau}{4}|H|$ such that the intersection $\bigcap_{s \in Y} s^{-1}A_1/t$ is non-empty. Let z be any point in this intersection and put $r := zt$. Then Yt is a subset of both A_1 and Hr , which proves that $|A_1 \cap Hr| > \frac{\tau}{4}|H|$. \square

Proof of (F-B5). Let $E(\overline{\mathcal{G}})$ denote the collection of all idempotent ultrafilters that extend the filter \mathcal{G} . It follows from Lemma 2.1 of [30] that if a set is contained in every ultrafilter $p \in E(\overline{\mathcal{G}})$, then the set is \mathcal{G} -syndetic. Let $p \in E(\overline{\mathcal{G}})$ and $R \in p$ be arbitrary. Hence, to prove (F-B5) it suffices to show that $\{t \in \mathbf{S} : d^*(A \cap t^{-1}A; \mathcal{H}, \{\mathbf{S}\}) > 0\} \cap R \neq \emptyset$.

Choose $\delta > 0$ such that $d^*(A; \mathcal{F}, \{\mathbf{S}\}) \geq \delta$ and let $\tau \in (0, \delta)$ be arbitrary. Thus, we can find $W \in \mathcal{F}$ such that for all $F \in \mathcal{P}_f(W)$ there exists $t \in \mathbf{S}$ such that $|A \cap Ft| > \tau|F|$. Fix any number $r \in \mathbb{N}$ with $r \geq \frac{8}{\tau}$ and define $\varepsilon := \frac{\tau^2}{16}$. Then, using property (FC), we can find $V \in \mathcal{G}$ such that for all $H \in \mathcal{P}_f(W)$ we can find a non-empty finite subset $F \subset W$ such that $|F \triangle sF| < \varepsilon|F|$ for all $s \in H$. Note that $V \cap R \in p$. Any set contained in an idempotent ultrafilter contains (arbitrarily large) finite-sum-sets [26, Theorem 5.12]. This means that we can find numbers $x_1, \dots, x_r \in \mathbf{S}$ such that the set

$$\Gamma := \{x_{i_1}x_{i_2} \cdots x_{i_k} : 1 \leq i_1 < i_2 < \dots < i_k \leq r, k \leq r\}$$

is contained in $V \cap R$. Now, we make the claim that

$$d^*\left(A \cap \bigcup_{x \in \Gamma} x^{-1}A; \mathcal{G}, \{G\}\right) \geq \frac{\tau^2}{16}. \quad (5.2)$$

Indeed, if we can verify this claim then, using part (F-B4), we deduce that there exists some $x \in \Gamma$ such that $d^*(A \cap x^{-1}A; \mathcal{H}, \{G\}) > 0$ and this shows that $\{x \in \mathbf{S} : d^*(A \cap x^{-1}A; \mathcal{H}, \{\mathbf{S}\}) > 0\} \cap R \neq \emptyset$.

Let $H \in \mathcal{P}_f(V)$ be arbitrary. Then, since the sets Γ and H are contained in V , we can find a non-empty finite subset F of W such that $|F \triangle sF| < \varepsilon|F|$ for all $s \in H \cup \Gamma$. Then, since $F \subset W$, we can find $t \in \mathbf{G}$ such that $|A \cap Ft| > \tau|F|$. Using $|F \triangle sF| < \varepsilon|F|$ for all $x \in \Gamma$, we get that

$$|x^{-1}A/t \cap F| > |A \cap xFt| > (\tau - \varepsilon)|F| > \frac{\tau}{2}|F|.$$

If we set $\eta := \frac{\tau}{2}$ and $A_x := x^{-1}A/t = \{s \in \mathbf{S} : xst \in A\}$, then we get $|A_x \cap F| > \eta|F|$ for all $x \in \{x_1, x_1x_2, x_1x_2x_3, \dots, x_1 \cdots x_r\} \subset \Gamma$. Also, note that $r > 1 + \frac{\eta - \eta^2}{\varepsilon}$. So, using Lemma A.3, we can find two distinct elements y and z in the set $\{x_1, x_1x_2, x_1x_2x_3, \dots, x_1 \cdots x_r\}$ such that

$$|A_y \cap A_z \cap F| > (\eta^2 - \varepsilon)|F|. \quad (5.3)$$

Suppose $y = x_1 \cdots x_i$ and $z = x_1 \cdots x_j$ and let us assume without loss of generality that $i < j$. Then the element $x := y^{-1}z$ is given by $x_{i+1} \cdots x_j$ and is therefore contained in Γ . Also, note that $\eta^2 - \varepsilon = \frac{3\tau^2}{16}$. So, we conclude from (5.3) that

$$|A/t \cap x^{-1}A/t \cap zF| > \frac{3\tau^2}{16}|F|.$$

Then, using that $|F \triangle zF| < \varepsilon|F|$ and $|F \triangle sF| < \varepsilon|F|$ for all $s \in H$ we get that

$$|A/t \cap x^{-1}A/t \cap sF| > \frac{\tau^2}{16}|F|, \quad \forall s \in H.$$

Then, by applying Lemma A.2 to the sequence $(s^{-1}(A/t \cap x^{-1}A/t))_{s \in H}$, we can find $Y \subset H$ with $|Y| > \frac{\tau^2}{16}|H|$ such that the intersection $\bigcap_{s \in Y} s^{-1}(A/t \cap x^{-1}A/t)$ is non-empty. Let a be any point in this intersection and put $r := at$. Then Yr is a subset of both A and $x^{-1}A$. This shows that $|A \cap x^{-1}A \cap Hr| > \frac{\tau^2}{16}|F|$. Since H was chosen arbitrarily, this proves (5.2). \square

6. Connections between Theorem F and the Density Hales-Jewett Theorem

In this section we prove that Theorem F and the Density Hales-Jewett Theorem are equivalent.

The Density Hales-Jewett Theorem is, as the name suggests, a density version of the Hales-Jewett Theorem. For its statement recall the definition of variable words and combinatorial lines given in Section 3.

Density Hales-Jewett Theorem ([20, 21]). *Suppose $A \subset \bigcup_{n \in \mathbb{N}} [k]^n$ satisfies*

$$\limsup_{n \in \mathbb{N}} \frac{|A \cap [k]^n|}{k^n} > 0.$$

Then A contains a combinatorial line.

To show that Theorem F implies the Density Hales-Jewett Theorem we need the following definition.

Definition 6.1 (cf. [20, 21]). A set $A \subset \bigcup_{n \in \mathbb{N}} [k]^n$ is called **stationary** if for all $l_1, l_2 \in \mathbb{N} \cup \{0\}$ and for all words $v \in [k]^{l_1}$ and $w \in [k]^{l_2}$ we have

$$\lim_{n \rightarrow \infty} \left| \frac{|A \cap [k]^n|}{k^n} - \frac{|A \cap v \frown [k]^{n-l} \frown w|}{k^{n-l}} \right| = 0,$$

where $l = l_1 + l_2$. Roughly speaking, this means that the number of ‘large’ words in A that start with v and end with w is independent of the choice of v and w .

Let $\text{DHJ}(k, \delta)$ refer to the statement that there exists $n \geq 1$ such that all sets $A \subset [k]^n$ with $|A| > \delta k^n$ must contain a combinatorial line. Define

$$\delta_{\min} := \inf\{\delta \in (0, 1] : \text{DHJ}(k, \delta) \text{ holds}\}. \quad (6.1)$$

It is well known that the Density Hales-Jewett Theorem is equivalent to $\delta_{\min} = 0$ (for details see [21, Section 2]).

Lemma 6.2. *There exists a stationary set $A \subset \bigcup_{n \in \mathbb{N}} [k]^n$ satisfying $\lim_{n \rightarrow \infty} |A \cap [k]^n|/k^n = \delta_{\min}$ and containing no combinatorial lines.*

Note, under the assumption that $\delta_{\min} = 0$ the statement of Lemma 6.2 becomes trivial, since one can take A to be the empty set. Hence, the only purpose of this lemma is to deal with the hypothetical case $\delta_{\min} > 0$ when one attempts to prove the Density Hales-Jewett Theorem by contradiction.

Proof of Lemma 6.2. It follows from the definition of δ_{\min} that DHJ(k, τ) does not hold for all $\tau \in (0, \delta_{\min})$. In other words for every such τ and every $n \geq 1$ there exists a set $A_{\tau, n}$ with $|A_{\tau, n} \cap [k]^n| \geq \tau k^n$ containing no combinatorial line. For $n \in \mathbb{N}$ put $\tau_n := \delta_{\min} - \frac{1}{n}$ and define $A_n := A_{\tau_n, n}$. We claim that

$$\lim_{n \rightarrow \infty} \frac{|A_n|}{k^n} = \delta_{\min}. \quad (6.2)$$

Indeed, this can be seen as follows: on the one hand $\frac{|A_n|}{k^n} \geq \tau_n$ and $\tau_n \rightarrow \delta_{\min}$, whereas on the other hand the set A_n does not contain a combinatorial line and therefore, using the definition of δ_{\min} , it follows that $\frac{|A_n|}{k^n} \leq \delta_{\min}$ for sufficiently large n .

Let A be defined as $A := \bigcup_{n \in \mathbb{N}} A_n$. From (6.2) we deduce that

$$\lim_{n \rightarrow \infty} \frac{|A \cap [k]^n|}{k^n} = \delta_{\min}.$$

Also, A contains no combinatorial lines because A_n contains none for all $n \geq 1$. We claim that A is stationary. Once we have verified this claim, the proof is completed.

Let $l_1, l_2 \in \mathbb{N}$, $v \in [k]^{l_1}$ and $w \in [k]^{l_2}$ be arbitrary. Define ${}_v A_w$ to be the set of all words in A that start with v and end with w . It then suffices to show that

$$\lim_{n \rightarrow \infty} \frac{|{}_v A_w \cap [k]^n|}{k^{n-l}} = \delta_{\min}. \quad (6.3)$$

Note that for sufficiently large n the ratio $|{}_v A_w \cap [k]^n|/k^{n-l}$ cannot exceed δ_{\min} because A , and therefore also ${}_v A_w$, contains no combinatorial lines. Hence, it remains to show that the density of ${}_v A_w$ is not smaller than δ_{\min} . However, for $M := k^l$, where $l = l_1 + l_2$, and for $n \geq l$ the sets $v \wedge [k]^{n-l} \wedge w$, as v and w run through all words in $[k]^{l_1}$ and $[k]^{l_2}$ respectively, partition the set

$[k]^n$ into M cells of equal size. Also, for all $\varepsilon > 0$ and for all sufficiently large n we have that $|A \cap [k]^n|/k^n > \delta_{\min} - \varepsilon$ as well as $|A \cap v \cap [k]^{n-l} \cap w|/k^{n-l} = |{}_v A_w \cap [k]^n|/k^{n-l} < \delta_{\min} + \varepsilon$. Therefore, it follows directly from Lemma A.4 that for all $\tau < \delta_{\min}$ one has

$$\frac{|{}_v A_w \cap [k]^n|}{k^{n-l}} > \tau$$

for all sufficiently large n . Indeed, since v and w were chosen arbitrarily, this proves that A is stationary. \square

Lemma 6.3. *Assume $k \geq 2$. Let the sequence (x_j) , the filter \mathcal{F} and the map $\varphi : \bigcup_{n \in \mathbb{N}} [k]^n \rightarrow \mathbb{N}$ all be as in the proof of Proposition 3.1. If $A \subset \bigcup_{n \in \mathbb{N}} [k]^n$ is stationary with $\lim_{n \rightarrow \infty} |A \cap [k]^n|/k^n = \delta$ then the \mathcal{F} -density of $\varphi(A_2)$ is at least δ , where A_2 denotes the set of all words in A that end with the letter 2.*

Proof. Let $V_{k,N}$ be as in (3.1). To show that $d^*(A_2; \mathcal{F}) \geq \delta$ we must show that for all $\tau \in [0, \delta)$, for all finite nonempty subsets $F \subset V_{k,1}$ and for all $N \in \mathbb{N}$ there exists $s \in V_{k,N}$ such that $|\varphi(A_2) \cap (F + s)| > \tau|F|$. Hence let $\tau \in [0, \delta)$, $F \in \mathcal{P}_f(V_{k,1})$ and $N \in \mathbb{N}$ be arbitrary.

First, for every $x \in F$ pick a word a_x such that $\varphi(a_x) = x$. Note that adding the letter 1 at the end of a word w does not change the image of w under the map φ . In particular, by adding multiple 1s at the end of the words a_x if necessary we can assume without loss of generality that all the words a_x have the same length, which we denote by l , and that $l \geq N$.

For $n \geq l$ let $B_{x,n}$ denote the collection of all words $w \in [k]^{n-l-1}$ such that $a_x \frown w \frown 2 \in A_2$. Since A is stationary, it follows that

$$\lim_{n \rightarrow \infty} \frac{|B_{x,n}|}{k^{n-l}} = \delta, \quad \forall x \in F.$$

So, there exists at least one n such that $|B_{x,n}| > \tau k^{n-l-1}$ for all $x \in F$. This allows us to apply Lemma A.2 to the sequence $(B_{x,n})_{x \in F}$. Therefore we can find a set $Y \subset F$ with $|Y| > \tau|F|$ such that $\bigcap_{x \in Y} B_{x,n} \neq \emptyset$. Let \tilde{v} be any element in this intersection and define $v := \tilde{v} \frown 2$. Now, since $\varphi(a_x \frown v) = \varphi(a_x) + \varphi(v)$ and since $a_x \frown v \in A_2$ for all $x \in Y$, we deduce that $(Y + s) \subset \varphi(A_2)$, where $s = \varphi(v)$. So,

$$|\varphi(A_2) \cap (F + s)| > |\varphi(A_2) \cap (Y + s)| = |Y| > \tau|F|.$$

This completes the proof. \square

Proposition 6.4. *Theorem F implies the Density Hales-Jewett Theorem.*

Proof. Assume $k \geq 2$. For the sake of a contradiction, let us assume that the Density Hales-Jewett Theorem is not true, i.e., let us assume that $\delta_{\min} > 0$ where δ_{\min} is defined as in (6.1). Let the sequence (x_j) , the filter \mathcal{F} and the map $\varphi : \bigcup_{n \in \mathbb{N}} [k]^n \rightarrow \mathbb{N}$ be as in the proof of Proposition 3.1. According to Lemma 6.2 there exists a stationary set $A \subset \bigcup_{n \in \mathbb{N}} [k]^n$ containing no combinatorial lines and such that $\lim_{n \rightarrow \infty} |A \cap [k]^n|/k^n = \delta_{\min}$. Then, using Lemma 6.3, it follows that $\varphi(A_2)$ has \mathcal{F} -density equal to δ_{\min} , where A_2 denotes the collection of all words in A that end in the letter 2. In particular, the \mathcal{F} -density of $\varphi(A_2)$ is positive. For $1 \leq i \leq k$, let $x_i : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbb{N}$ be given by $x_i(\alpha) = (i-1) \sum_{j \in \alpha} x_j$. Recall, these IP-sets are \mathcal{F} -measurable (as is shown in the proof of Proposition 3.1). Therefore, we can apply Theorem F and we obtain $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbb{N}$ such that $\{a + x_1(\alpha), \dots, a + x_k(\alpha)\} \subset \varphi(A_2)$.

Finally, since $a + x_i(\alpha) \in V_{k,1}$ can be written as $\sum_{j=1}^m e_j x_j$ for some $m \geq 1$, where $e_j = i$ whenever $j \in \alpha$ and where $e_m = 2$, it follows from the definition of φ that the pre-image of the k -term arithmetic progression $a + x_1(\alpha), \dots, a + x_k(\alpha)$ under φ is a combinatorial line contained in A_2 . This contradicts the fact that A was assumed to contain no combinatorial lines. \square

The remainder of this section is dedicated to proving that the converse of Proposition 6.4 holds as well. In other words, our next goal is to show that Theorem F implies the Density Hales-Jewett Theorem. We need the following definitions.

Definition 6.5. If $x, y : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ are IP-sets in a semigroup (\mathbf{S}, \cdot) then we say that y is an **IP-subset** of x if there are $\alpha_1 < \alpha_2 < \dots \in \mathcal{P}_f(\mathbb{N})$ such that $x(\alpha_j) = y(\{j\})$ for all $j \in \mathbb{N}$.

If $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ is a collection of IP-sets in \mathbf{S} then we say that $\{x_1, \dots, x_k\}$ is **non-collapsing under taking IP-subsets** if any IP-subset of x_i is not an IP-subset of x_j for all $i, j \in [k]$ with $i \neq j$.

Lemma 6.6. *Let $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ be IP-sets in a cancellative commutative semigroup $(\mathbf{S}, +)$. Then, after reordering x_1, \dots, x_k if necessary, one can find $l \in [k]$ and $\alpha_1 < \alpha_2 < \dots \in \mathcal{P}_f(\mathbb{N})$ such that if $y_i(\beta) := \sum_{j \in \beta} x_i(\alpha_j)$, $1 \leq i \leq k$, then*

- every IP-set in $\{y_{l+1}, \dots, y_k\}$ coincides with an IP-set in $\{y_1, \dots, y_l\}$;
- $\{y_1, \dots, y_l\}$ is non-collapsing under taking IP-subsets.

Proof. We prove this lemma by induction on k . For $k = 1$ there is nothing to show, hence suppose the claim has already been proven for $k-1$. This means we can find $l \in [k]$ and IP-subsets y_1, \dots, y_{k-1} of x_1, \dots, x_{k-1} such that $\{y_{l+1}, \dots, y_{k-1}\} \subset \{y_1, \dots, y_l\}$ and such that $\{y_1, \dots, y_l\}$ is non-collapsing under taking IP-subsets. Let y_k denote the corresponding IP-subset of x_k .

We distinguish two simple cases. If $\{y_1, \dots, y_l\} \cup \{y_k\}$ is non-collapsing under taking IP-subsets, then we are done immediately. If $\{y_1, \dots, y_l\} \cup \{y_k\}$ collapses when passing to some IP-subset, then we obtain a new collection of IP-subsets z_1, \dots, z_k such that $\{z_{l+1}, \dots, z_{k-1}\} \cup \{z_k\} \subset \{z_1, \dots, z_l\}$ and such that $\{z_1, \dots, z_l\}$ is non-collapsing under taking IP-subsets. This finishes the proof. \square

Lemma 6.7. *Suppose $y_1, \dots, y_l : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ is a collection of IP-sets in a cancellative commutative semigroup $(\mathbf{S}, +)$ that is non-collapsing under taking IP-subsets. Then for any finite set $\Omega \subset \mathbf{G}$ and any $\beta \in \mathcal{P}_f(\mathbb{N})$ one can find $\alpha \in \mathcal{P}_f(\mathbb{N})$ with $\alpha > \beta$ such that*

$$(\Omega + y_i(\alpha)) \cap (\Omega + y_j(\alpha)) = \emptyset$$

for all $i, j \in [k]$ with $i \neq j$.

Proof. Let $i_1, i_2 \in [k]$ with $i_1 \neq i_2$ be arbitrary. Suppose there exists some $\beta \in \mathcal{P}_f(\mathbb{N})$ such that for all $\alpha > \beta$ we have

$$(\Omega + y_{i_1}(\alpha)) \cap (\Omega + y_{i_2}(\alpha)) \neq \emptyset. \quad (6.4)$$

Define $y_\Delta := y_{i_1} - y_{i_2}$ and note that (6.4) is equivalent to

$$y_\Delta(\alpha) \in \Omega - \Omega \quad (6.5)$$

for all $\alpha > \beta$. This means we can find $c \in \Omega - \Omega$ and $j_1, j_2, \dots \in \mathbb{N}$ with $\beta < \{j_1\} < \{j_2\} < \dots$ such that $y_\Delta(j_1) = y_\Delta(j_2) = \dots = c$. However, this implies that $nc \in \Omega - \Omega$ for all $n \in \mathbb{N}$, because of (6.5) and the fact that $y_\Delta(j_1 + j_2 + \dots + j_n) = c + \dots + c = nc$. Since $\Omega - \Omega$ is a finite set, it follows that $mc = 0$ for some $1 \leq m \leq n$. Define $\alpha_i := \{j_{im+1}, j_{im+2}, \dots, j_{(i+1)m}\}$. Then $y_\Delta(\alpha_i) = 0$ for all $i \in \mathbb{N}$. This contradicts the fact that $i_1 \neq i_2$ and that the IP-sets y_{i_1} and y_{i_2} are non-collapsing under taking IP-subsets. \square

Lemma 6.8. *Assume \mathcal{F} is an idempotent filter on a cancellative commutative semigroup $(\mathbf{S}, +)$ and $y_1, \dots, y_l : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ is a collection of \mathcal{F} -measurable IP-sets that is non-collapsing under taking IP-subsets. Then for*

every $V \in \mathcal{F}$ there exist $\alpha_1 < \alpha_2 < \dots \in \mathcal{P}_f(\mathbb{N})$ and $v_1, v_2, \dots \in \mathbf{S}$ such that if

$$\Omega_n := \left\{ \sum_{j=1}^n v_j + y_{i_j}(\alpha_j) : i \in [l] \right\},$$

then $\Omega_n \subset V$ and $|\Omega_n| = l^n$.

Proof. Let $V \in \mathcal{F}$ be given. In order to prove the lemma we will prove the following slightly stronger statement by induction: There are $\alpha_1 < \alpha_2 < \dots \in \mathcal{P}_f(\mathbb{N})$ and $v_1, v_2, \dots \in \mathbf{G}$ such that for all $n \geq 1$ there exists $V_n \in \mathcal{F}$ with $0 \in V_n$ and such that

$$V_n + \Omega_n = V_n + \left\{ \sum_{j=1}^n v_j + y_{i_j}(\alpha_j) : i \in [l] \right\} \subset V$$

and $|\Omega_n| = l^n$. We proceed by induction on n . Suppose we have already found $V_n \in \mathcal{F}$, $v_1, \dots, v_n \in \mathbf{S}$ and $\alpha_1 < \dots < \alpha_n \in \mathcal{P}_f(\mathbb{N})$. Consider the set $V_n - \mathcal{F} = \{s \in \mathbf{S} : -s + V \in \mathcal{F}\}$ and note that since \mathcal{F} is idempotent the set $V_n - \mathcal{F}$ is contained in \mathcal{F} . Also, $0 \in V_n - \mathcal{F}$. Since all the IP-sets y_i , $i \in [l]$, are \mathcal{F} -measurable there exist \mathbb{E} -many $\alpha \in \mathcal{P}_f(\mathbb{N})$ such that $y_i(\alpha) \in V_n - \mathcal{F}$. Hence, using Lemma 6.7, we can find $\alpha_{n+1} \in \mathcal{P}_f(\mathbb{N})$ such that $\alpha_n < \alpha_{n+1}$, such that $y_i(\alpha) \in V_n - \mathcal{F}$ holds for all $i \in [l]$ and such that

$$(\Omega_n + y_i(\alpha_{n+1})) \cap (\Omega_n + y_j(\alpha_{n+1})) = \emptyset \quad (6.6)$$

for all $i, j \in [l]$ with $i \neq j$. This implies that

$$V'_{n+1} := \bigcap_{i=1}^k V_n - y_i(\alpha_{n+1})$$

is contained in \mathcal{F} . Finally, let v_{n+1} be any element in the intersection $(V'_{n+1} - \mathcal{F}) \cap V'_{n+1}$ and define

$$V_{n+1} := V'_{n+1} - v_{n+1}.$$

Clearly, $V_{n+1} \in \mathcal{F}$ and $0 \in V_{n+1}$. Also, $V_{n+1} + v_{n+1} + y_i(\alpha_{n+1}) \in V_n$ for all $i \in [l]$. From this it follows that

$$V_{n+1} + \Omega_{n+1} \subset V.$$

Moreover, (6.6) shows that $|\Omega_{n+1}| = l|\Omega_n| = l^{n+1}$. This completes the inductive proof. \square

Proposition 6.9. *The Density Hales-Jewett Theorem implies Theorem F.*

Proof. Let $k \in \mathbb{N}$, let \mathcal{F} be an idempotent filter on a cancellative commutative semigroup $(\mathbf{S}, +)$ and let $x_1, \dots, x_k : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{S}$ be \mathcal{F} -measurable IP-sets. Suppose $A \subset \mathbf{S}$ with $d^*(A; \mathcal{F}) \geq \delta$ is given. We want to find elements $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{S}$ such that $\{a + x_1(\alpha), \dots, a + x_k(\alpha)\} \subset A$.

First, we invoke Lemma 6.6. This means that after switching to IP-subsets of x_1, \dots, x_k if necessary we can assume without loss of generality that $\{x_{l+1}, \dots, x_k\} \subset \{x_1, \dots, x_l\}$ and that x_1, \dots, x_l is non-collapsing under taking IP-subsets. In particular, it suffices to show that we can find $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{S}$ such that $\{a + x_1(\alpha), \dots, a + x_l(\alpha)\} \subset A$.

Using the definition of upper \mathcal{F} -density, we can find $V \in \mathcal{F}$ such that for all $\tau \in [0, \delta)$ and all $F \in \mathcal{P}_f(V)$ there exists some s such that $|A \cap (F + s)| > \tau|F|$. Then, using Lemma 6.8, we can find $\alpha_1 < \alpha_2 < \dots \in \mathcal{P}_f(\mathbb{N})$ and $v_1, v_2, \dots \in \mathbf{S}$ such that

$$\Omega_n := \left\{ \sum_{j=1}^n v_j + x_{i_j}(\alpha_j) : i \in [l] \right\} \subset V,$$

and $|\Omega_n| = l^n$. For every n the set Ω_n is a finite subset of V , hence there exists some s_n such that $|A \cap (\Omega_n + s_n)| > \tau|\Omega_n|$. We define the map $\xi_n : [l]^n \rightarrow (\Omega_n + s_n)$ as

$$\xi_n(a_1 a_2 \cdots a_n) := s_n + \sum_{j=1}^n v_j + x_{a_j}(\alpha_j).$$

This map is a bijection, because $|\Omega_n| = l^n$. Let B denote the set $B := \bigcup_{n \in \mathbb{N}} \xi_n^{-1}(A \cap (s_n + \Omega_n))$. We observe that $B \subset \bigcup_{n \in \mathbb{N}} [l]^n$ and that

$$\limsup_{n \in \mathbb{N}} \frac{|B \cap [l]^n|}{l^n} > \tau.$$

Hence, it follows from the Density Hales-Jewett Theorem that B contains at least one combinatorial line. In other words, there exists a variable word $w(*)$ such that $\{w(1), \dots, w(l)\} \subset B$. Let n denote the length of $w(*)$. We can find $\alpha \in \mathcal{P}_f(\mathbb{N})$ and $a \in \mathbf{S}$ such that $\xi_n(\{w(1), \dots, w(l)\}) = \{a + x_1(\alpha), \dots, a + x_l(\alpha)\}$. Finally, since $\xi_n(\{w(1), \dots, w(l)\}) \subset (\Omega_n + s_n) \cap A$ we conclude that $\{a + x_1(\alpha), \dots, a + x_l(\alpha)\} \subset A$. \square

7. A Proof of Theorem D

We begin this section by discussing various properties of polynomial mappings as defined in Definition 1.3. It is not hard to see that if $P : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ is a polynomial mapping then so is P^{-1} . Also, it is shown in [7] that for nilpotent groups \mathbf{G} the product PQ of two polynomial mappings $P, Q : \mathcal{P}_f(\mathbb{N}) \rightarrow \mathbf{G}$ is itself a polynomial mapping.

We are particularly interested in collections of \mathcal{F} measurable polynomial mappings for idempotent filters \mathcal{F} on a nilpotent group \mathbf{G} . Let us give an example of such a setup.

Example 7.1. Let (x_n) and (y_n) be sequences of positive integers. Let \mathbf{G} denote the discrete Heisenberg group, i.e.,

$$\mathbf{G} = \left\{ \begin{pmatrix} 1 & a & c \\ & 1 & b \\ & & 1 \end{pmatrix} : a, b, c \in \mathbb{Z} \right\},$$

and recall that \mathbf{G} is 2-step nilpotent. Let

$$V_N := \left\{ \begin{pmatrix} 1 & \sum_{N \leq i \leq M} a_i x_i & \sum_{N \leq \max\{i,j\} \leq M} c_{ij} x_i y_j \\ & 1 & \sum_{N \leq j \leq M} b_j y_j \\ & & 1 \end{pmatrix} : \begin{array}{l} M \geq N, \\ a_i, b_j \in \{0, 1\}, \\ c_{ij} \in \{-1, 0, 1\} \end{array} \right\}.$$

Then the filter $\mathcal{F} = \{V \subset \mathbf{G} : \exists N \in \mathbb{N} \text{ s.t. } V_N \subset V\}$ is an idempotent filter on \mathbf{G} , as can be checked by straightforward calculations. For $\alpha \in \mathcal{P}_f(\mathbb{N})$ define

$$x_\alpha := \begin{pmatrix} 1 & \sum_{i \in \alpha} x_i & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \quad y_\alpha := \begin{pmatrix} 1 & 0 & 0 \\ & 1 & \sum_{j \in \alpha} y_j \\ & & 1 \end{pmatrix} \quad (7.1)$$

and

$$z_{\alpha \times \beta} := \begin{pmatrix} 1 & 0 & \sum_{(i,j) \in \alpha \times \beta} x_i y_j \\ & 1 & 0 \\ & & 1 \end{pmatrix}. \quad (7.2)$$

On the one hand, for any fixed $\beta \in \mathcal{P}_f(\mathbb{N})$ the maps $\alpha \mapsto x_\alpha$, $\alpha \mapsto y_\alpha$, $\alpha \mapsto z_{\alpha \times \beta}$ and $\alpha \mapsto z_{\beta \times \alpha}$ are \mathcal{F} measurable polynomial mappings of degree 1. On the other hand, the maps $\alpha \mapsto x_\alpha y_\alpha$, $\alpha \mapsto y_\alpha x_\alpha$ and $\alpha \mapsto z_{\alpha \times \alpha}$ are \mathcal{F} -measurable polynomial mappings of degree 2.

In the following, we refer to any finite collection of polynomial mappings $\mathcal{A} = \{P_1, \dots, P_m\}$ as a **system**.

Proposition 7.2 ([7]). *Suppose (\mathbf{G}, \cdot) is a nilpotent group. There exists a set \mathcal{W} , called the **set of weights**, endowed with a linear ordering \prec , such that for every system \mathcal{A} there is an element $w(\mathcal{A}) \in \mathcal{W}$ associated to \mathcal{A} , referred to as the **weight** of \mathcal{A} , having the following properties:*

- (i) *For every system \mathcal{A} there exist \mathbb{E} -many $\beta \in \mathcal{P}_f(\mathbb{N})$ such that $w(\mathcal{A}) = w(\mathcal{A} \cup \Delta_\beta \mathcal{A})$, where $\Delta_\beta P$ denotes the ‘directional derivative’*

$$\Delta_\beta P(\alpha) := P(\alpha \cup \beta)P^{-1}(\alpha);$$

- (ii) *Let $P_{\text{low}} \in \mathcal{A}$ and let $\mathcal{A}' = \{PP_{\text{low}}^{-1} : P \in \mathcal{A}\}$. If $w(\{P_{\text{low}}\}) \leq w(\{P\})$ for all $P \in \mathcal{A}$ and P_{low} does not equal $1_{\mathbf{G}}$ \mathbb{E} -a.e. then $w(\mathcal{A}') < w(\mathcal{A})$;*
- (iii) *The set \mathcal{W} contains a minimal element $e_{\mathcal{W}}$ and $w(\mathcal{A}) = e_{\mathcal{W}}$ if and only if $P = 1_{\mathbf{G}}$ \mathbb{E} -a.e. for all $P \in \mathcal{A}$;*
- (iv) *If $c_1, \dots, c_m \in \mathbf{G}$, if \mathcal{A} is a system and if $\mathcal{A}' = \bigcup_{i=1}^m \mathcal{A}^{c_i} = \bigcup_{i=1}^m c_i^{-1} \mathcal{A} c_i$ then $w(\mathcal{A}) = w(\mathcal{A}')$.*

Definition 7.3 (**P**-minimal systems). Let \mathbf{P} be a good collection of \mathcal{F} -measurable polynomial mappings for an idempotent filter \mathcal{F} on a nilpotent group (\mathbf{G}, \cdot) . We say that a system \mathcal{A} is a **P-minimal system** with **P-minimal element** P_{\min} if P_{\min} is contained in \mathcal{A} and if $\mathcal{A}P_{\min}^{-1} = \{PP_{\min}^{-1} : P \in \mathcal{A}\}$ is a subset of \mathbf{P} .

Lemma 7.4. *Let \mathbf{P} be a good collection of \mathcal{F} -measurable polynomial mappings for an idempotent filter \mathcal{F} on a nilpotent group (\mathbf{G}, \cdot) . Suppose \mathcal{A} is a **P-minimal system** with **P-minimal element** P_{\min} . Then there exists $C \in \mathcal{F}$ such that for all finite non-empty subsets $G \subset C$ the system*

$$\bigcup_{c \in G} \mathcal{A}^c = \{P^c : P \in \mathcal{A}, c \in G\}$$

*is **P**-minimal with **P**-minimal element P_{\min} .*

Proof. First, note that every element $P \in \mathcal{A}$ can be written as RP_{\min} for some $R \in \mathbf{P}$. Then, using the definition of good collections of \mathcal{F} -measurable polynomial mappings, for every such R we can find an \mathcal{F} -large set C_R such that $R^c[c, P_{\min}^{-1}] \in \mathbf{P}$ for all $c \in C_R$. Let C denote the intersection $\bigcap_R C_R$. Note that the set C is \mathcal{F} -large, as it is an intersection of finitely many \mathcal{F} -large sets.

To finish the proof it suffices to show that for every $P \in \mathcal{A}$ and for every $c \in C$ the polynomial mapping $P^c P_{\min}^{-1}$ lies in \mathbf{P} . However, simple

algebra manipulations show that $P^c P_{\min}^{-1} = R^c[c, P_{\min}^{-1}]$ and this proves the claim. \square

Lemma 7.5. *Let \mathbf{P} be a good collection of \mathcal{F} -measurable polynomial mappings for an idempotent filter \mathcal{F} on a nilpotent group (\mathbf{G}, \cdot) . Suppose \mathcal{A} is a \mathbf{P} -minimal systems with \mathbf{P} -minimal element P_{\min} containing the constant polynomial $1_{\mathbf{G}}$. Then there exists $C \in \mathcal{F}$ and $\beta_0 \in \mathcal{P}_f(\mathbb{N})$ such that if there exist $u \in C$ and $\beta > \beta_0$ such that $P(\beta)u \in C$ for all $P \in \mathcal{A}$ then*

$$\mathcal{A}' := \mathcal{A}^u \cup \Delta_{\beta} \mathcal{A}^u$$

is \mathbf{P} -minimal with \mathbf{P} -minimal element $P_{\min}^{\text{new}}(\alpha)$, which we can take to be $\Delta_{\beta} P_{\min}^u$.

Proof. First, note that every element $P \in \mathcal{A}$ can be written as RP_{\min} for some $R \in \mathbf{P}$. Define $c := P_{\min}(\beta)u$. On the one hand we have

$$\begin{aligned} \Delta_{\beta} P^u (P_{\min}^{\text{new}})^{-1} &= c^{-1} P_{\min}(\beta) P^{-1}(\beta) P(\alpha \cup \beta) P_{\min}^{-1}(\beta \cup \alpha) c \\ &= c^{-1} \Delta_{\beta} (P P_{\min}^{-1}) c \\ &= \Delta_{\beta} R^c \\ &= (RD_{\beta} R)^c \end{aligned}$$

On the other hand, we have

$$\begin{aligned} P^u (P_{\min}^{\text{new}})^{-1} &= (u^{-1} P(\alpha) u) (u^{-1} P_{\min}^{-1}(\alpha \cup \beta) P_{\min}(\beta) u) \\ &= u^{-1} R(\alpha) P_{\min}(\alpha) P_{\min}^{-1}(\beta \cup \alpha) P_{\min}(\beta) u \\ &= (RD_{\beta} P_{\min}^{-1})^u \end{aligned}$$

Now the claim follows directly from the definition of good collections of \mathcal{F} -measurable polynomial mappings. \square

For the proof of Theorem D we will use PET-induction, a technique that was developed in [5], and which proceeds by induction on the weight of systems as defined in Proposition 7.2. However, this inductive process requires us to replace Theorem D with the stronger Theorem H below, so that at every inductive step we are able to rely on a strong enough induction hypothesis.

Theorem H. *Let \mathbf{P} be a good collection of \mathcal{F} -measurable polynomial mappings for an idempotent filter \mathcal{F} on a nilpotent group (\mathbf{G}, \cdot) . Let \mathcal{A} be a \mathbf{P} -minimal systems with \mathbf{P} -minimal element P_{\min} , partitioned into three classes, $\mathcal{A} = \mathcal{A}^- \cup \{1_{\mathbf{G}}\} \cup \mathcal{A}^+$. Then for all $\beta \in \mathcal{P}_f(\mathbb{N})$, for all $V \in \mathcal{F}$ and*

for all piecewise \mathcal{F} -syndetic sets A there exist $\alpha \in \mathcal{P}_f(\mathbb{N})$ with $\alpha > \beta$ and $v \in V$ such that

$$\begin{aligned} P(\alpha)v &\in V, & \forall P \in \mathcal{A}^-, \\ P(\alpha)v &\in A, & \forall P \in \{1_{\mathbf{G}}\} \cup \mathcal{A}^+. \end{aligned}$$

Theorem D is indeed an immediate consequence of Theorem H, because one can choose \mathcal{A}^- to be the empty set.

The following lemma will be instrumental in proving Theorem H.

Lemma 7.6 (Color Focusing). *Let \mathbf{P} be a good collection of \mathcal{F} -measurable polynomial mappings for an idempotent filter \mathcal{F} on a nilpotent group (\mathbf{G}, \cdot) . Suppose $\mathcal{A} = \mathcal{A}^- \cup \{1_{\mathbf{G}}\} \cup \mathcal{A}^+$ is a given \mathbf{P} -minimal system with \mathbf{P} -minimal element P_{\min} and assume that Theorem H has already been proven for all \mathbf{P} -minimal systems $\mathcal{B} = \mathcal{B}^- \cup \{1_{\mathbf{G}}\} \cup \mathcal{B}^+$ with $w(\mathcal{B}^+) < w(\mathcal{A}^+)$. Let G be a finite non-empty subset of \mathbf{G} such that $\mathcal{A}_0 = \bigcup_{c \in G} \mathcal{A}^c$ is also a \mathcal{F} -minimal system with the same \mathcal{F} -minimal element P_{\min} . Assume B is a subset of \mathbf{G} such that $G^{-1}B \in \mathcal{F}$. Then, for every $s \geq 1$, one of the following two cases holds:*

- (1) *For all $U \in \mathcal{F}$ and $\alpha_0 \in \mathcal{P}_f(\mathbb{N})$ there exist $u_s \in U$, $\alpha_1, \dots, \alpha_s \in \mathcal{P}_f(\mathbb{N})$ with $\alpha_0 < \alpha_1 < \dots < \alpha_s \in \mathcal{P}_f(\mathbb{N})$ and s distinct elements $c_1, \dots, c_s \in G$ such that*

$$P(\alpha_j \cup \dots \cup \alpha_s)u_s \in U, \quad \forall P \in \mathcal{A}_0^- \cup \{1_{\mathbf{G}}\}, \quad (7.3)$$

$$P(\alpha_j \cup \dots \cup \alpha_s)u_s \in c_j^{-1}B, \quad \forall P \in \mathcal{A}_0^+. \quad (7.4)$$

Moreover the system $\mathcal{A}_s = \mathcal{A}_0^{u_s} \cup \Delta_{\alpha_s} \mathcal{A}_0^{u_s} \cup \dots \cup \Delta_{\alpha_1 \cup \dots \cup \alpha_s} \mathcal{A}_0^{u_s}$ remains \mathbf{P} -minimal with \mathbf{P} -minimal element $\Delta_{\alpha_1 \cup \dots \cup \alpha_s} P_{\min}^{u_s}$.

- (2) *For all $U \in \mathcal{F}$ and $\alpha_0 \in \mathcal{P}_f(\mathbb{N})$ there exist $\alpha \in \mathcal{P}_f(\mathbb{N})$ with $\alpha > \alpha_0$, $c \in G$ and $v \in cU$, such that*

$$\begin{aligned} P(\alpha)v &\in cU, & \forall P \in \mathcal{A}^-, \\ P(\alpha)v &\in B, & \forall P \in \{1_{\mathbf{G}}\} \cup \mathcal{A}^+. \end{aligned}$$

Proof. In this proof it will be convenient to identify polynomial mappings with their equivalency class of \mathbb{E} -a.e. equivalent polynomial mappings. This is allowed since the statement of Lemma 7.6 as well as all proceeding arguments in this proof are insensitive to replacing polynomial mappings with elements in their \mathbb{E} -a.e. equivalence class.

We proceed by induction on s and start with $s = 1$. We can write \mathcal{A}_0 as $\mathcal{A}_0^- \cup \{1_{\mathbf{G}}\} \cup \mathcal{A}_0^+$, where $\mathcal{A}_0^{\pm} = \bigcup_{c \in G} c^{-1} \mathcal{A}^{\pm} c$. We know that $w(\mathcal{A}^+) =$

$w(\mathcal{A}_0^+)$, because of Proposition 7.2, part (iv). Since we only care about \mathbb{E} -a.e. equivalency classes we tacitly assume that in the decomposition $\mathcal{A}_0 = \mathcal{A}_0^- \cup \{1_G\} \cup \mathcal{A}_0^+$ all polynomial mappings in \mathcal{A}_0 that are \mathbb{E} -a.e. equal to 1_G are grouped with $\{1_G\}$ and that \mathcal{A}_0^+ and \mathcal{A}_0^- contain no polynomial mapping that is \mathbb{E} -a.e. equal to 1_G . Let us pick $P_{\text{low}} \in \mathcal{A}_0^+$ such that $w(\{P_{\text{low}}\}) \leq w(\{P\})$ for all $P \in \mathcal{A}_0^+$. Define

$$\mathcal{B}_0^- = \{PP_{\text{low}}^{-1} : P \in \mathcal{A}_0^- \cup \{1_G\}\} \text{ and } \mathcal{B}_0^+ = \{PP_{\text{low}}^{-1} : P \in \mathcal{A}_0^+ \setminus \{P_{\text{low}}\}\}.$$

It follows from Proposition 7.2, part (ii), that $w(\mathcal{B}_0^+) < w(\mathcal{A}_0^+)$. Also, after a moment's consideration, we see that $\mathcal{B}_0 = \mathcal{B}_0^- \cup \{1_G\} \cup \mathcal{B}_0^+$ is an \mathcal{F} -minimal system with \mathcal{F} -minimal element $Q_{\min} = P_{\min}P_{\text{low}}^{-1}$. Using Lemma 7.5 we can find an \mathcal{F} -large set $U_0 \subset U$ and $\beta_0 \in \mathcal{P}_f(\mathbb{N})$ with $\beta_0 > \alpha_0$ such that $\mathcal{A}_0^{u_1} \cup \Delta_{\alpha_1}\mathcal{A}^{u_1}$ is \mathcal{F} -minimal with \mathcal{F} -minimal element $\Delta_{\alpha_1}P_{\min}^{u_1}$ whenever $\alpha_1 > \beta_0$ and $P(\alpha_1)u_1 \in U_0$ for all $P \in \mathcal{A}_0$.

Since $G^{-1}B \in \mathcal{F}$, there exists some $c_1 \in G$ such that $c_1^{-1}B$ is piecewise \mathcal{F} -syndetic (cf. Remark 2.4). Let us put $N_0 := c_1^{-1}B \cap U_0$ and let us note that the set N_0 is a piecewise \mathcal{F} -syndetic set, as it is an intersection of a piecewise \mathcal{F} -syndetic set and an \mathcal{F} -large set. Since the weight of the system \mathcal{B}_0^+ is strictly smaller than the weight of \mathcal{A}^+ , we can apply Theorem H to β_0 , U_0 , N_0 and to the system \mathcal{B}_0 in order to find $\alpha_1 > \beta_0$ and $u'_1 \in U_0$ such that

$$\begin{aligned} Q(\alpha_1)u'_1 &\in U_0, & \forall Q \in \mathcal{B}_0^-, \\ Q(\alpha_1)u'_1 &\in N_0, & \forall Q \in \{1_G\} \cup \mathcal{B}_0^+. \end{aligned}$$

If we put $u_1 = P_{\text{low}}^{-1}(\alpha_1)u'_1$, then a simple calculation shows that $Q(\alpha_1)u'_1 = P(\alpha_1)u_1$. So with this choice of u_1 , α_1 and c_1 equations (7.3) and (7.4) are satisfied and the system $\mathcal{A}_0^{u_1} \cup \Delta_{\alpha_1}\mathcal{A}^{u_1}$ remains \mathcal{F} -minimal. This concludes the case $s = 1$.

Next, let us deal with the inductive step, $s \rightarrow s+1$. Take any ultrafilter $q \in K(\overline{\mathcal{F}})$ with $B \in q$ (cf. Theorem 2.3). Let $B' := B/q$ and let $U' := (U/\mathcal{F}) \cap ((G^{-1}B)/\mathcal{F})$. Observe that B' is piecewise \mathcal{F} -syndetic, by virtue of Theorem 2.2, and we see that $U' \in \mathcal{F}$. Also, since $G^{-1}B \in \mathcal{F}$, it follows that $(G^{-1}B)/q \in \mathcal{F}$. Then, a simple calculation shows that $G^{-1}(B/q) = (G^{-1}B)/q$, which tells us that $G^{-1}B' \in \mathcal{F}$.

This means we can apply the induction hypothesis to U' and B' in order to find $u_s \in U'$, elements $\alpha_1, \dots, \alpha_s \in \mathcal{P}_f(\mathbb{N})$ with $\alpha_0 < \alpha_1 < \dots < \alpha_s$, and s distinct 'colors' $c_1, \dots, c_s \in G$ such that

$$\begin{aligned} P(\alpha_j \cup \dots \cup \alpha_r)u_s &\in U', & \forall P \in \mathcal{A}_0^- \cup \{1_G\}, \\ P(\alpha_j \cup \dots \cup \alpha_r)u_s &\in c_j^{-1}B', & \forall P \in \mathcal{A}_0^+. \end{aligned}$$

This implies that there exist sets $B'' \in q$ and $U'' \in \mathcal{F}$ such that

$$P(\alpha_j \cup \dots \cup \alpha_s)u_s U'' \subset U, \quad \forall P \in \mathcal{A}_0^- \cup \{1_G\}, \quad (7.5)$$

$$P(\alpha_j \cup \dots \cup \alpha_s)u_s B'' \subset c_j^{-1}B, \quad \forall P \in \mathcal{A}_0^+. \quad (7.6)$$

Since $u_s \in U/\mathcal{F}$ we may assume that $u_s U'' \subset U$, because otherwise we can replace U'' with $U'' \cap u_s^{-1}U$. Analogously, since $u_s \in (G^{-1}B)/\mathcal{F}$ we may assume that $u_s B'' \subset G^{-1}B$ because otherwise we can replace B'' with $B'' \cap u_s^{-1}G^{-1}B$. Then, since $G^{-1}B$ covers $u_s B''$, there exists a piecewise \mathcal{F} -syndetic subset $N_s \subset B''$ and a ‘color’ $c_{s+1} \in F$ such that $u_s N_s \subset c_{s+1}^{-1}B$.

Now we have to distinguish two cases. The first case is when $c_{r+1} = c_j$ for some $j \in \{1, \dots, r\}$. In this case one may take any $n \in U'' \cap N_s$ and put $c := c_{r+1}$, $v := cu_s n$ and $\alpha := \alpha_j \cup \dots \cup \alpha_r$. With this choice of $\alpha \in \mathcal{P}_f(\mathbb{N})$, $c \in G$ and $v \in cU$ we are in case (2) of Lemma 7.6 and therefore we are done with the current inductive step $s \rightarrow s+1$, as well as with all subsequent inductive steps, and the inductive process terminates here.

The second case is when $c_{s+1} \neq c_j$ for all $j \in \{1, \dots, s\}$. If this is the case then we proceed as follows. Define

$$\mathcal{A}_s^\pm := u_s^{-1}(\mathcal{A}_0^\pm \cup \Delta_{\alpha_s} \mathcal{A}_0^\pm \cup \dots \cup \Delta_{\alpha_1 \cup \dots \cup \alpha_s} \mathcal{A}_0^\pm)u_s.$$

Under the assumptions of the induction hypothesis the system \mathcal{A}_s is \mathcal{F} -minimal with \mathcal{F} -minimal element $\Delta_{\alpha_1 \cup \dots \cup \alpha_s} P_{\min}^{u_s}$. Using Lemma 7.5 we can find a \mathcal{F} -large set $U_s \subset U''$ and $\beta_s \in \mathcal{P}_f(\mathbb{N})$ with $\beta_s > \alpha_s$ such that $\mathcal{A}_s^u \cup \Delta_{\alpha_{s+1}} \mathcal{A}_s^u$ is \mathcal{F} -minimal with \mathcal{F} -minimal element $\Delta_{\alpha_1 \cup \dots \cup \alpha_{s+1}} P_{\min}^{u_s u}$ whenever $\alpha_{s+1} > \beta_s$ and $P(\alpha_{s+1})u \in U_s$ for all $P \in \mathcal{A}_s$. Let us put $N_s := B'' \cap U_s$. Since B'' is contained in q and since U_s is contained in \mathcal{F} it follows that N_s is contained in q . Therefore, using Theorem 2.3, we deduce that N_s is piecewise \mathcal{F} -syndetic.

Let P_{low} denote the element of \mathcal{A}_s^+ of lowest weight, i.e. $w(\{P_{\text{low}}\}) \leq w(\{P\})$ for all $P \in \mathcal{A}_s^+$. Again, we assume without loss of generality that \mathcal{A}_s^+ contains no polynomial mappings that are \mathbb{E} -a.e. equivalent to 1_G . In particular, P_{low} is not \mathbb{E} -a.e. equal to 1_G . Define

$$\mathcal{B}_s^- = \{PP_{\text{low}}^{-1} : P \in \mathcal{A}_s^- \cup \{1_G\}\} \text{ and } \mathcal{B}_s^+ = \{PP_{\text{low}}^{-1} : P \in \mathcal{A}_s^+ \setminus \{P_{\text{low}}\}\}.$$

We have $w(\mathcal{B}_s^+) < w(\mathcal{A}_s^+)$ by Proposition 7.2, part (ii). Also, $\mathcal{B}_s = \mathcal{B}_s^- \cup \{1_G\} \cup \mathcal{B}_s^+$ is an \mathcal{F} -minimal system with \mathcal{F} -minimal element $Q_{\min} = \Delta_{\alpha_1 \cup \dots \cup \alpha_s} P_{\min}^u P_{\text{low}}^{-1}$. This means we can apply Theorem H to β_s , U_s , N_s and to the system \mathcal{B}_s to find $\alpha_{s+1} > \beta_s$ and $u'_{s+1} \in U_s$ such that

$$Q(\alpha_{s+1})u'_{s+1} \in U_s, \quad \forall Q \in \mathcal{B}_s^-, \quad (7.7)$$

$$Q(\alpha_{s+1})u'_{s+1} \in N_s, \quad \forall Q \in \{1_G\} \cup \mathcal{B}_s^+. \quad (7.8)$$

Finally, let us define $u_{s+1} := u_s P_{\text{low}}^{-1}(\alpha_{s+1}) u'_{s+1}$. For every $P \in \mathcal{A}_0$ and every $j \in [s]$ we can find $Q \in \mathcal{B}_s$ such that

$$\begin{aligned} Q(\alpha_{s+1}) u'_{s+1} &= u_s^{-1} P^{-1}(\alpha_j \cup \dots \cup \alpha_s) P(\alpha_j \cup \dots \cup \alpha_{s+1}) u_s P_{\text{low}}^{-1}(\alpha_{s+1}) u'_{s+1} \\ &= u_s^{-1} P^{-1}(\alpha_j \cup \dots \cup \alpha_s) P(\alpha_j \cup \dots \cup \alpha_{s+1}) u_{s+1}. \end{aligned}$$

Hence, if we combine equations (7.7) and (7.8) with equations (7.5) and (7.6), then we get

$$\begin{aligned} P(\alpha_j \cup \dots \cup \alpha_{s+1}) u_{s+1} &\subset U, & \forall P \in \mathcal{A}_0^- \cup \{1_{\mathbf{G}}\}, \\ P(\alpha_j \cup \dots \cup \alpha_{s+1}) u_{s+1} &\subset c_j^{-1} B, & \forall P \in \mathcal{A}_0^+ \end{aligned}$$

for all $j \in [s]$. For the case $j = s+1$ we simply note that $u_s U_S \subset u_s U'' \subset U$ and that $u_s N_s \subset c_{s+1}^{-1} B$ and therefore it follows from equations (7.7) and (7.8) and the fact that $\mathcal{A}_0^{u_s} \subset \mathcal{A}_s$ that

$$\begin{aligned} P(\alpha_{s+1}) u_{s+1} &\subset U, & \forall P \in \mathcal{A}_0^- \cup \{1_{\mathbf{G}}\}, \\ P(\alpha_{s+1}) u_{s+1} &\subset c_{s+1}^{-1} B, & \forall P \in \mathcal{A}_0^+. \end{aligned}$$

Finally, the newly created system \mathcal{A}_{s+1} is \mathcal{F} -minimal because $\alpha_{s+1} > \beta_s$. This completes the inductive step $s \rightarrow s+1$. \square

Proof of Theorem H. We proceed by induction on the weight of \mathcal{A}^+ . The beginning of the induction is for $\mathcal{A}^+ = \emptyset$. In this case we have to find $\alpha > \beta$ and $v \in A$ such that $P(\alpha)v \in V$ for all $P \in \mathcal{A}^-$. Let P_{\min} denote the \mathcal{F} -minimal element of $\mathcal{A} = \mathcal{A}^- \cup \{1_{\mathbf{G}}\}$. Since P_{\min}^{-1} is in $\mathcal{A} P_{\min}^{-1}$ and since all the elements in $\mathcal{A} P_{\min}^{-1}$ are \mathcal{F} -measurable, we conclude that P_{\min}^{-1} is \mathcal{F} -measurable.

The set $\{v \in V : v^{-1}A \text{ is piecewise } \mathcal{F}\text{-syndetic}\}$ is \mathcal{F} -large and therefore the set $W = \{\alpha \in \mathcal{P}_f(\mathbb{N}) : P_{\min}(\alpha)A \text{ is piecewise } \mathcal{F}\text{-syndetic}\}$ is \mathbb{E} -large. Moreover PP_{\min}^{-1} is \mathcal{F} -measurable for all $P \in \mathcal{A}$ and therefore there exists some $\alpha \in W$ such that $(PP_{\min}^{-1})(\alpha) \in V/\mathcal{F}$ for all $P \in \mathcal{A}^-$. In other words the set $(P_{\min}P^{-1})(\alpha)V$ is contained in \mathcal{F} for every $P \in \mathcal{A}^-$. Hence the following intersection is also contained in \mathcal{F} ,

$$V' = \bigcap_{P \in \mathcal{A}^-} P_{\min}(\alpha) P^{-1}(\alpha) V \in \mathcal{F}.$$

Since the intersection of a piecewise \mathcal{F} -syndetic set with an \mathcal{F} -large set is always non-empty, it follows that we can find some $v' \in V' \cap P_{\min}(\alpha)A$. Now put $v = P_{\min}^{-1}(\alpha)v'$. Then clearly, $v \in A$ and $P(\alpha)v \in V$ for all $P \in \mathcal{A}^-$. This completes the initial step of the induction.

For the inductive step assume that Theorem H has already been proven for all systems $\mathcal{B}^- \cup \{1_{\mathbf{G}}\} \cup \mathcal{B}^+$ with $w(\mathcal{B}^+) < w(\mathcal{A}^+)$. Let $C \in \mathcal{F}$ be as guaranteed by Lemma 7.4. Fix any ultrafilter $q \in K(\overline{\mathcal{F}})$ with $A \in q$ and let A' denote the syndetic set A/p (cf. Theorem 2.2).

Since A' is \mathcal{F} -syndetic, by definition, we can find a finite non-empty set $G \subset V/\mathcal{F} \cap C$ such that $G^{-1}A' \in \mathcal{F}$. We apply Lemma 7.6 with $s = |G| + 1$, $U = \bigcap_{c \in G} c^{-1}V/p$, and to $B = A'$. Since $s > |G|$ we can not be in the case (1) of Lemma 7.6 and therefore we have to be in case (2). This means we can find $\alpha > \beta$, $c \in G$ and $v' \in U$ such that

$$\begin{aligned} P(\alpha)v' \in cU \subset V/p, & \quad \forall P \in \mathcal{A}^-, \\ P(\alpha)v' \in A' = A/p, & \quad \forall P \in \{1_G\} \cup \mathcal{A}^+. \end{aligned}$$

So we can find a p -large set N such that

$$\begin{aligned} P(\alpha)v'N \in V, & \quad \forall P \in \mathcal{A}^-, \\ P(\alpha)v'N \in A, & \quad \forall P \in \{1_G\} \cup \mathcal{A}^+. \end{aligned}$$

Hence, we can choose v to be any element in $v'N$ and the proof is completed. \square

We end this section with a proof of Theorem E

Proof of Theorem E. Fix $k \geq 1$. Let $\mathbb{Z}[x]$ denote the collection of all polynomials with integer coefficients, let $S : \mathbb{Z}[x] \rightarrow \mathbb{Z}[x]$ denote the map $S(p(x)) = p(x+1)$ and let $R : \mathbb{Z}[x] \rightarrow \mathbb{Z}[x]$ denote the map $R(p(x)) = p(x) + x^k$. Let \mathbf{G} denote the group generated by S and R . It is well known and straightforward to check that \mathbf{G} is a k -step nilpotent group.

We start with the following claim: Let $\Sigma_{\leq k+1}$ denote the collection of all words in the letters $*_1$ and $*_2$ in which each variable $*_1$ and $*_2$ occurs at most k times. Then $w_1(R, S) \neq w_2(R, S)$ for all $w_1, w_2 \in \Sigma_{\leq k+1}$ with $w_1(*_1, *_2) \neq w_2(*_1, *_2)$.

To prove this claim let $w \in \Sigma_{\leq k}$ be arbitrary. We can write

$$w(R, S) = S^{v_0} R^{u_1} S^{v_1} R^{u_2} S^{v_2} \dots R^{u_\ell} S^{v_\ell} R^{u_{\ell+1}}$$

with the conditions

- $u_{\ell+1}, v_0 \in [k] \cup \{0\}$;
- $u_j, v_j \in [k]$ for all $j \geq 1$;
- $\sum_{j=0}^{\ell} v_j \leq k$;
- $\sum_{j=1}^{\ell+1} u_j \leq k$.

Put $m_j = v_0 + v_1 + \dots + v_{j-1}$. Then $w(R, S)$ applied to the polynomial x^k yields

$$w(R, S)x^k = u_1(x + m_1)^k + \dots + u_{\ell+1}(x + m_{\ell+1})^k$$

Observe that $\{(x + m)^k : m \in [k] \cup \{0\}\}$ forms a linearly independent subset of $\mathbb{Z}[x]$. Hence if $w_1, w_2 \in \Sigma_{\leq k+1}$ with

$$w_1(R, S)x^k = u_{1,1}(x + m_{1,1})^k + \dots + u_{1,\ell+1}(x + m_{1,\ell+1})^k$$

and

$$w_2(R, S)x^k = u_{2,1}(x + m_{2,1})^k + \dots + u_{2,\ell+1}(x + m_{2,\ell+1})^k$$

then $w_1(R, S) = w_2(R, S)$ if and only if $u_{1,j} = u_{2,j}$ and $m_{1,j} = m_{2,j}$ for all j . However, from $m_{1,j} = m_{2,j}$ it follows that $v_{1,j} = v_{2,j}$ and therefore $w_1(*_1, *_2) = w_2(*_1, *_2)$. In other words, this shows that one can find non-degenerated nilprogressions of length k and rank 2 in \mathbf{G} .

On the other hand, one can show that \mathbf{G} does not admit non-degenerated nilprogressions of length $(k+1)$ and rank 2. This follows from the fact that $\{(x + m)^k : m \in [k+1] \cup \{0\}\}$ does not form a linearly independent subset of $\mathbb{Z}[x]$.

Next, let $(\mathbf{G}_n)_{n \in \mathbb{N}}$ be \mathbb{N} -many identical copies of \mathbf{G} , let R_n and S_n denote identical copies of the maps R and S and suppose G_n is generated by R_n and S_n . Let $\mathbf{G}_\infty := \bigoplus_{n \in \mathbb{N}} \mathbf{G}_n$. For convenience we shall identify \mathbf{G}_n with its embedding into \mathbf{G}_∞ . In particular we view R_n and S_n as elements in \mathbf{G}_∞ in the obvious way.

For $\alpha \in \mathcal{P}_f(\mathbb{N})$ define

$$S_\alpha := \prod_{n \in \alpha} S_n \quad \text{and} \quad R_\alpha := \prod_{n \in \alpha} R_n.$$

For $\gamma \in \mathcal{P}_f(\mathbb{N})$ let

$$U_\gamma := \{S_\alpha R_\beta : \alpha, \beta > \gamma\}$$

and let

$$\mathcal{F} := \{U \subset \mathbf{G}_\infty : \exists \gamma \text{ such that } U_\gamma[\mathbf{G}_\infty, \mathbf{G}_\infty] \subset U\}.$$

Note that \mathcal{F} is an idempotent filter. This is easy to see if one interprets \mathcal{F} as the pull-back of an idempotent filter on the abelian group $\mathbf{G}_\infty/[\mathbf{G}_\infty, \mathbf{G}_\infty]$ under the natural quotient map.

Let $V := U_{\{1\}}$. We make the following claim, which will finish the proof of Theorem E: For any partition of V into finitely many classes, some class contains a non-degenerated nilprogression of length k and rank 2.

Define

$$\mathbf{P} := \{\alpha \mapsto w(R_\alpha, S_\alpha) : w \in \Sigma_{\leq k+1}\}.$$

Then \mathbf{P} is a good collection of \mathcal{F} -measurable polynomial mappings, as can be shown by a (rather bothersome) calculation. Therefore, using Theorem D, for every partitioning of V into finitely many classes we can find $a \in \mathbf{G}_\infty$ and $\alpha \in \mathcal{P}_f(\mathbb{N})$ such that

$$\{aw(R_\alpha, S_\alpha) : w \in \Sigma_{\leq k+1}\}$$

is contained in one single class. However, the set $\{w(R_\alpha, S_\alpha) : w \in \Sigma_{\leq k+1}\}$ is in a 1-1 correspondence with the set $\{w(R, S) : w \in \Sigma_{\leq k+1}\}$ and hence $w_1(R_\alpha, S_\alpha) \neq w_2(R_\alpha, S_\alpha)$ for all $w_1, w_2 \in \Sigma_{\leq k+1}$ with $w_1(*_1, *_2) \neq w_2(*_1, *_2)$. In particular, this means that the nilprogression $\{aw(R_\alpha, S_\alpha) : w \in \Sigma_{\leq k+1}\}$ is non-degenerated, which finishes the proof. \square

A. Appendix

The following result is due to V. Bergelson.

Theorem A.1 ([1, Theorem 1.1]). *Let (X, \mathcal{B}, μ) be a probability space and let $\varepsilon > 0$. If $A_1, A_2, \dots \in \mathcal{B}$ satisfy*

$$\mu(A_i) > \varepsilon,$$

then there exists $D \subset \mathbb{N}$ with $d(D) > \varepsilon$ and such that for all $H \in \mathcal{P}_f(D)$,

$$\mu\left(\bigcap_{i \in H} A_i\right) > 0.$$

The above theorem possesses the following finitistic version:

Lemma A.2. *Let X be a finite set, let $\tau > 0$ and let $(A_x)_{x \in F}$ be a sequence of subsets of X satisfying*

$$|A_x| > \tau|X|, \quad \text{for all } x \in F.$$

Then there exists $Y \subset F$ with $|Y| > \tau|F|$ and such that $\bigcap_{y \in Y} A_y \neq \emptyset$.

Proof. Assume there exists a finite set X and sets $A_x \subset X$, $x \in F$ with $|A_x| > \tau|X|$ and such that $\bigcap_{y \in Y} A_y = \emptyset$ for all $Y \subset F$ with $|Y| > \tau|F|$. Without loss of generality we can assume that $F = \{1, 2, \dots, M\}$ for some $M \in \mathbb{N}$. Let μ be the normalized counting measure on X . From the finite sequence A_1, \dots, A_M we can construct an infinite sequence $A_1, A_2, \dots \subset X$

by following the rule $A_i = A_j$ if $i \equiv j \pmod{M}$. Then clearly, any set $D \subset \mathbb{N}$ that satisfies

$$\bigcap_{i \in H} A_i \neq \emptyset, \quad \forall H \in \mathcal{P}_f(D)$$

also has $\bar{d}(D) \leq \tau$. But this contradicts the statement of Theorem A.1. \square

Lemma A.3. *Let F be a finite set and let $\eta, \varepsilon \in (0, 1)$. Suppose X is a finite set with $|X| > 1 + \frac{\eta - \eta^2}{\varepsilon}$ and suppose for all $x \in X$ the set A_x is a subset of F satisfying*

$$|A_x| > \eta|F|.$$

Then there are $x, y \in X$, $x \neq y$, such that $|A_x \cap A_y| > (\eta^2 - \varepsilon)|F|$.

Lemma A.4. *For every $\tau, \delta_{\inf} \in (0, 1]$ with $\tau < \delta_{\inf}$ and any $M \in \mathbb{N}$ there exists $\varepsilon = \varepsilon(\tau, \delta_{\inf}, M) > 0$ such that the following holds: Suppose $m \leq M$ and suppose $X = X_1 \cup \dots \cup X_m$ is a partitioning of a finite set X into m cells of equal size, i.e. $|X_i| = |X_j|$ for all $i, j \in [m]$. Then for all $A \subset X$, if*

$$\frac{|A|}{|X|} > \delta_{\inf} - \varepsilon, \tag{A.1}$$

then either

$$\frac{|A \cap X_j|}{|X_j|} > \delta_{\inf} + \varepsilon, \quad \text{for some } j \in [m],$$

or

$$\frac{|A \cap X_i|}{|X_i|} > \tau, \quad \forall i \in [m].$$

Proof. Suppose $|A \cap X_j| \leq |X_j|(\delta_{\inf} + \varepsilon)$ for all $j \in [m]$. Clearly

$$\frac{|A|}{|X|} = \frac{1}{m} \sum_{i \in [m]} \frac{|A \cap X_i|}{|X_i|}.$$

Let $i \in [m]$ be arbitrary. From the above and from (A.1) we deduce that

$$m(\delta_{\inf} - \varepsilon) < \frac{|A \cap X_i|}{|X_i|} + (m-1)(\delta_{\inf} + \varepsilon)$$

and so

$$\delta_{\inf} - (m-1)\varepsilon < \frac{|A \cap X_i|}{|X_i|}.$$

Hence if ε is chosen sufficiently small then $\tau < \delta_{\inf} - (m-1)\varepsilon$. \square

References

- [1] V. BERGELSON, *Sets of recurrence of \mathbf{Z}^m -actions and properties of sets of differences in \mathbf{Z}^m* , J. London Math. Soc. (2), 31 (1985), pp. 295–304.
- [2] ———, *Ergodic Ramsey theory—an update*, in Ergodic theory of \mathbf{Z}^d actions (Warwick, 1993–1994), vol. 228 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1996, pp. 1–61.
- [3] ———, *Ultrafilters, IP sets, dynamics, and combinatorial number theory*, in Ultrafilters across mathematics, vol. 530 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2010, pp. 23–47.
- [4] V. BERGELSON AND N. HINDMAN, *Ultrafilters and multidimensional Ramsey theorems*, Combinatorica, 9 (1989), pp. 1–7.
- [5] V. BERGELSON AND A. LEIBMAN, *Polynomial extensions of van der Waerden’s and Szemerédi’s theorems*, J. Amer. Math. Soc., 9 (1996), pp. 725–753.
- [6] ———, *Set-polynomials and polynomial extension of the Hales-Jewett theorem*, Ann. of Math. (2), 150 (1999), pp. 33–75.
- [7] ———, *Topological multiple recurrence for polynomial configurations in nilpotent groups*, Adv. Math., 175 (2003), pp. 271–296.
- [8] V. BERGELSON AND R. MCCUTCHEON, *An ergodic IP polynomial Szemerédi theorem*, Mem. Amer. Math. Soc., 146 (2000), pp. viii+106.
- [9] V. BERGELSON AND R. MCCUTCHEON, *Idempotent ultrafilters, multiple weak mixing and Szemerédi’s theorem for generalized polynomials*, J. Anal. Math., 111 (2010), pp. 77–130.
- [10] V. BERGELSON AND I. RUZSA, *Squarefree numbers, IP sets and ergodic theory*, in Paul Erdős and his mathematics, I (Budapest, 1999), vol. 11 of Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 2002, pp. 147–160.
- [11] E. BREUILLARD AND B. GREEN, *Approximate groups. I: The torsion-free nilpotent case*, J. Inst. Math. Jussieu, 10 (2011), pp. 37–57.
- [12] E. BREUILLARD, B. GREEN, AND T. TAO, *The structure of approximate groups*, Publ. Math. Inst. Hautes Études Sci., 116 (2012), pp. 115–221.
- [13] ———, *Small doubling in groups*, in Erdős centennial, vol. 25 of Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 2013, pp. 129–151.
- [14] E. BREUILLARD AND M. TOINTON, *Nilprogressions and groups with moderate growth*, Adv. Math., 289 (2016), pp. 1008–1055.
- [15] D. DAVENPORT, *The minimal ideal of compact subsemigroups of βS* , Semi-group Forum, 41 (1990), pp. 201–213.
- [16] D. DE, N. HINDMAN, AND D. STRAUSS, *Sets central with respect to certain subsemigroups of βS_d* , Topology Proc., 33 (2009), pp. 55–79.

- [17] R. ELLIS, *Distal transformation groups*, Pacific J. Math., 8 (1958), pp. 401–405.
- [18] H. FURSTENBERG, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures.
- [19] H. FURSTENBERG AND Y. KATZNELSON, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. Analyse Math., 45 (1985), pp. 117–168.
- [20] ———, *A density version of the Hales-Jewett theorem for $k = 3$* , Discrete Math., 75 (1989), pp. 227–241. Graph theory and combinatorics (Cambridge, 1988).
- [21] ———, *A density version of the Hales-Jewett theorem*, J. Anal. Math., 57 (1991), pp. 64–119.
- [22] H. FURSTENBERG AND B. WEISS, *Topological dynamics and combinatorial number theory*, J. Analyse Math., 34 (1978), pp. 61–85 (1979).
- [23] J. T. GRIESMER, *Recurrence, rigidity, and popular differences*, ArXiv e-prints, (2015).
- [24] A. W. HALES AND R. I. JEWETT, *Regularity and positional games*, Trans. Amer. Math. Soc., 106 (1963), pp. 222–229.
- [25] N. HINDMAN AND D. STRAUSS, *Density in arbitrary semigroups*, Semigroup Forum, 73 (2006), pp. 273–300.
- [26] ———, *Algebra in the Stone-Čech compactification*, de Gruyter Textbook, Walter de Gruyter & Co., Berlin, 2012. Theory and applications, Second revised and extended edition [of MR1642231].
- [27] S. KAKEYA AND S. MORIMOTO, *On a theorem of MM. Bandet and van der Waerden*, Japanese Journal of Mathematics, 7 (1930), pp. 163–165.
- [28] J. NĚSTŘIL AND V. RÖDL, *unpublished*.
- [29] K. NUMAKURA, *On bicomact semigroups*, Math. J. Okayama Univ., 1 (1952), pp. 99–108.
- [30] O. SHUUNGULA, Y. ZELENYUK, AND Y. ZELENYUK, *The closure of the smallest ideal of an ultrafilter semigroup*, Semigroup Forum, 79 (2009), pp. 531–539.
- [31] J. SPENCER, *Restricted Ramsey configurations*, J. Combinatorial Theory Ser. A, 19 (1975), pp. 278–286.
- [32] A. TSERUNYAN, *A Ramsey theorem on semigroups and a general van der Corput lemma*, ArXiv e-prints, (2014).
- [33] B. L. VAN DER WAERDEN, *Beweis einer baudetschen vermutung*, Nieuw. Arch. Wisk., 15 (1928), pp. 212–216.

- [34] M. WALTERS, *Combinatorial proofs of the polynomial van der Waerden theorem and the polynomial Hales-Jewett theorem*, J. London Math. Soc. (2), 61 (2000), pp. 1–12.
- [35] P. ZORIN-KRANICH, *A nilpotent IP polynomial multiple recurrence theorem*, J. Anal. Math., 123 (2014), pp. 183–225.

John H. Johnson

DEPARTMENT OF MATHEMATICS, OHIO STATE UNIVERSITY, COLUMBUS, OH 43210, USA
johnson.5316@osu.edu

Florian K. Richter

DEPARTMENT OF MATHEMATICS, OHIO STATE UNIVERSITY, COLUMBUS, OH 43210, USA
richter.109@osu.edu